BMJ Open

Evaluation of the diagnostic accuracy of FeNO in patients with suspected asthma: study protocol for a prospective diagnostic study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-045420
Article Type:	Protocol
Date Submitted by the Author:	30-Sep-2020
Complete List of Authors:	Kellerer, Christina; Technical University of Munich, Institute of General Practice and Health Services Research Hapfelmeier, Alexander; Technical University of Munich, Institute of General Practice and Health Services Research Joerres, Rudolf; Ludwig-Maximilians-University Munich, Occupational and Environmental Medicine Schultz, Konrad; Bad Reichenhall Clinic, Centre for Rehabilitation, Pneumology and Orthopedics Brunn, Benjamin; Technical University of Munich, Institute of General Practice and Health Services Research Schneider, Antonius; Technical University Munich, Institute of General Practice
Keywords:	Asthma < THORACIC MEDICINE, Chronic airways disease < THORACIC MEDICINE, GENERAL MEDICINE (see Internal Medicine)

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Evaluation of the diagnostic accuracy of FeNO in patients with suspected asthma: study protocol for a prospective diagnostic study

Christina Kellerer[1]*, Alexander Hapfelmeier[1,2], Rudolf A. Jörres[3], Konrad Schultz[4], Benjamin Brunn[1], Antonius Schneider[1]

[1] Technical University of Munich, School of Medicine, Institute of General Practice and Health Services Research, Munich, Germany

[2] Institute of Medical Informatics, Statistics and Epidemiology, School of Medicine, Technical University of Munich, Munich, Germany

[3] Institute and Outpatient Clinic for Occupational, Social and Environmental Medicine, Ludwig-Maximilians-Universität München, Munich, Germany

[4] Clinic Bad Reichenhall, Center for Rehabilitation, Pneumology and Orthopedics, Bad Reichenhall

Corresponding Author:

Christina Kellerer, MSc

TUM School of Medicine

Institute of General Practice and Health Services Research

Technical University of Munich

Orleansstraße 47

81667 Munich, Germany

Mail: christina.kellerer@mri.tum.de

Phone: +49 89 6146589-17

Abstract

Introduction

The measurement of fractional exhaled nitric oxide (FeNO) is promising for diagnosing asthma and might substitute for bronchial provocation (BP) tests. To evaluate the diagnostic accuracy of FeNO within a confirmatory study, the following hypotheses will be tested: 1. A FeNO cut-off > 50ppb is suitable for diagnosing asthma (sensitivity 35%, specificity 95%). 2. If the clinical symptoms "allergic rhinitis" and "wheezing" are present, asthma can be diagnosed at FeNO > 33ppb with a positive predictive value (PPV) ≥ 70%. 3. A FeNO > 33ppb can predict responsiveness to inhaled corticosteroid (ICS) with a PPV ≥ 70%.

Methods and analysis

A prospective diagnostic study will be conducted in three practices of pneumologists in Germany. 300 patients suspected of suffering from asthma will be included. As an index test, patients perform FeNO measurement with the device NIOX VERO®. As reference a test, patients are examined with whole bodyplethysmography and BP, if necessary. After three months, patients with an asthma diagnosis will be examined again to verify the diagnosis and evaluate ICS responsiveness. Patients who did not receive an asthma diagnosis at the initial examination will be phoned after three months and asked about persistent respiratory symptoms to exclude false negative findings. As a primary target, sensitivity and specificity of FeNO > 50ppb will be determined. As a secondary target the PPV for asthma at FeNO > 33ppb, when the symptoms "allergic rhinitis" and "wheezing" are present, will be calculated. Regarding ICS responsiveness, the PPV of FeNO > 33ppb will be determined.

Ethics and dissemination

The study was approved by the Ethical Committee of the Technical University of Munich (Reference number 122/20 S). The major results will be published in peer-reviewed academic journals and disseminated through conferences.

Trial registration

German Clinical Trials Register (DRKS00021125).

Key words: FeNO, asthma, ICS responsiveness, bronchial provocation

Strengths and limitations of this study

- As this prospective confirmatory study aims to validate pre-defined FeNO cut-off values for an asthma diagnosis and ICS responsiveness it might be able to determine the appropriate place of FeNO in the diagnosis of asthma and in routine care.

- A high quality reference standard will be used in this study as the diagnosis of asthma will be made in all patients based on bronchial provocation tests assessed in whole body plethysmography and a potential asthma diagnosis will be verified after three months.
- FeNO devices from various manufacturers should be compared since it cannot be excluded that optimal cut-off values differ between devices.
- The present study is not able to assess the impact of FeNO on diagnostic decision making in routine care and patient outcomes.

1. Introduction

Background

The diagnosis of asthma is limited by the fact that airway obstruction is often not present during investigation by spirometry or whole body plethysmography (WBP) when patients suffer from mild symptoms, thus leading to diagnostic uncertainty. For these cases, diagnostic guidelines recommend bronchial provocation (BP) tests, which can only be performed in pneumologic centres in order to diagnose or exclude asthma [1, 2]. Moreover, peak-flow variability can be assessed, but the low diagnostic value of this method has been demonstrated and it is considered as a second choice method [3, 4]. Thus, in the case of inconclusive lung function results, BP remains the reference standard for the diagnosis of asthma [1, 2].

Numerous studies have demonstrated that, in addition to BP, the measurement of fractional exhaled nitric oxide (FeNO) has a high potential for diagnosing asthma and could possibly replace BP [5, 6]. Nitric oxide (NO) is released during type-2 allergic inflammation [7] and it could be shown that patients with asthma, even in mild stages of the disease, exhale NO in higher concentrations [8]. In contrast to BP, FeNO is a non-invasive measurement that can be performed without risk to the patient in a short time.

The available studies indicate that a cut-off value of 50ppb is well suited for diagnosing asthma [9, 10]. However, such values were identified only by post-hoc analyses in the sense of multiple and exploratory testing. Accordingly, the major criticism is that the diagnostic value of the cut-off points identified and proposed so far need to be confirmed in a prospective study [1, 9].

It was shown in a secondary analysis that even lower FeNO values than 50ppb could be useful for diagnosis when considering appropriate anamnestic information. If, for example, the patient suffers from allergic rhinitis and wheezing, an asthma diagnosis can be established with a high degree of certainty when FeNO is >33 ppb [11]. However, this algorithm needs to be validated in a multicentre study. Studies also indicate that the diagnostic accuracy of FeNO measurement might be superior to BP (e.g. [9, 11]), as the latter gives correctly positive values in only about 70% of cases [12]. This might be especially true for allergic, inflammatory

alterations of respiratory tract, which might be better diagnosed via FeNO than BP [13]. In line with this, FeNO could be suitable for predicting responsiveness to inhaled corticosteroide (ICS) in asthma. The study by Martin et al. [14] showed that FeNO > 33ppb could be used to predict the response to ICS in patients with suspected asthma with a high degree of certainty. However, these values were also identified by post hoc analyses. Another study found FeNO values ≥40ppb to predict ICS responsiveness in patients with non-specific respiratory symptoms [15]. In view of these reports, it is obvious that a prospective confirmatory study is necessary to validate pre-defined cut-off values and to determine the appropriate place of FeNO in the diagnosis of asthma as well as in routine care.

Aims of the study

The present study aims to evaluate sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) to clarify the following hypotheses:

1. Primary hypothesis: The sensitivity of FeNO measurement for diagnosing asthma is 35% at the cut-off > 50ppb, and specificity is 95%.
2. Secondary hypothesis: if the clinical symptoms "allergic rhinitis" and "wheezing" are present, the PPV of FeNO > 33ppb is at least 70% (validation of the diagnostic algorithm [10]). Sensitivity, specificity and NPV will also be estimated.
3. Further secondary hypothesis: The PPV of FeNO > 33ppb for ICS responsiveness is at least 70%. Sensitivity, specificity and NPV will also be estimated.

2. Methods and analysis

Trial design

The study will be conducted as a multi-centre diagnostic study in three practices of pneumologists. Patients with suspected asthma visiting one of the three practices will be informed about the study. After having obtained the informed consent of the patient, FeNO measurement will be carried out.

Afterwards, the patient will be routinely examined with WBP and, if necessary, BP to clarify a potential asthma diagnosis. This diagnostic procedure is routinely performed in German practices of pneumologists in ambulatory care if asthma is suspected.

Three months after inclusion, patients who have been diagnosed as suffering from asthma will be invited into the practices for a follow-up examination. Based on this examination with WBP and BP (when appropriate), it will be verified whether the patient has responded positively to ICS (delayed type of diagnostic study [16]). Based on the recommendations of the national [1] and international [2] asthma guidelines, this time interval is reasonable for a therapy of at least three months after initial diagnosis [1, 2]. Patients who did not receive an asthma diagnosis will be phoned after three months and asked whether the respiratory symptoms still persist in

order to rule out false negative findings. Patients with persistent symptoms will be invited back for re-evaluation.

Study setting

The study will be conducted in three private practices of pneumologists in Germany ("Zentrum für Pneumologie, Onkologie und Schlafmedizin am Diakonissenkrankenhaus" in Augsburg; "Lungenpraxis Starnberg" in Starnberg; "Pneumologie Elisenhof" in Munich). Further practices will be included if necessary for sufficient recruitment within the intended time frame.

Eligibility criteria

Inclusion criteria:

All adult patients suspected of suffering from asthma, who visit one of the three participating practices of pneumologists and declare their written informed consent to participate in the study will be included consecutively.

Exclusion criteria:

Patients with the following criteria are excluded:

- Patients who do not agree to participate in this study
- Patients younger than 18 years (legal grounds)
- Patients who do not understand the meaning of the study due to a lack of knowledge of the German language
- Patients with already diagnosed obstructive airway disease
- Patients who smoked on the day of the examination (distortion of the FeNO results and reactivity during BP testing)
- Nitrate-rich meal (e.g. salad) before the examination (false high FeNO values)
- Patients with respiratory infection < 6 weeks before examination (distortion of the FeNO results and/or BP)

Recruitment and taking informed consent

Patients visiting one of the three participating practices of pneumologists will be contacted by a doctoral candidate (BB) from the Medical Faculty of the Technical University of Munich or a Research Associate at the Institute of General Practice and Health Services Research of the Technical University of Munich regarding possible participation in the study. They will check the inclusion and exclusion criteria and will inform the patient about the study. Finally, the attending pneumologist will provide detailed information about the study. In the case of patients who are screened for participation but who ultimately do not participate in the study (due to disagreement or other reasons), age, gender, and the reason for non-participation will be

documented anonymously in order to be able to conduct a non-responder analysis to assess a potential recruitment bias.

Interventions

Patients included in the study will be examined at first contact (time point t1) and 3 months later (time point t2). The diagnostic work-up is summarized in figure 1 and the patient timeline in figure 2.

Initial examination (t1):

During the first presentation for diagnostic work-up in one of the participating practices of pneumologists, patients are examined with a NO-measuring device (NIOX VERO®) as an index test. Afterwards patients will be examined with WBP as a reference standard; a BP test is performed additionally as part of the diagnostic routine, if required by the pneumologist. In addition, patients complete a questionnaire with structured questions about medical history and symptoms. The questionnaire also contains the "Asthma Control Questionnaire (ACQ)" [17]. The ACQ is used to determine the extent of asthma control (controlled, partially controlled, uncontrolled) and the responsiveness to ICS. ICS responsiveness is given if the ACQ score improves by at least 0.5 in the sense of a "minimal important difference" [14, 18].

Index test:

The index test is performed with the electrochemically-based NO-measuring device NIOX VERO®. This device is CE-certified, available in national and international markets, and is already widely used in practices. The FeNO measurements are performed once for each patient according to the recommendations of the ATS and ERS [19]. It is a non-invasive measurement since the patient only needs to take a deep breath through the device and exhale evenly. The FeNO measurements are performed by a doctoral candidate, or a research assistant, or a lung function assistant according to the instructions of the manufacturer.

Reference test:

Following the FeNO measurement, an examination with WBP is routinely performed and, if required by the pneumologist, BP is performed as part of the diagnostic routine to rule-in or rule-out the diagnosis of asthma.

WBP**:** WBP is considered as the reference standard used to diagnose obstructive airway diseases. An obstructive airway disease is indicated if FEV_1 and/or FEV_1/FVC are below their lower limits of normal [20]. A reversible airway obstruction is diagnosed if the bronchodilation test is positive (ΔFEV_1 > 12% and 200ml). If there is no bronchial obstruction, BP is performed.

Bronchial provocation test (BP): BP is performed to determine bronchial hyperresponsiveness (BHR) to methacholine according to the 1-concentration-4-step dosimeter protocol [21]. This yields similar results as the ATS multi-concentration protocol [22] but offers advantages in clinical practice. The test is considered positive (indicating BHR) if FEV_1 decreases by at least

20% after inhalation of a maximum cumulative methacholine dose of 960µg, and/or if specific airway resistance (sRaw) increases simultaneously by at least 100% and to at least 2.0 kPa*s, and/or if airway resistance (Raw) increases simultaneously by at least 100% and to at least 0.5 kPa*s/L [22, 23].

Follow-up examination after 3 months (t2):

A single BP test as a reference standard for the diagnosis of asthma only reflects the situation at the time of examination. In some cases, patients with a positive BP test do not suffer from asthma (false positive), since the positive predictive value of BP is only about 70% [12, 24]. According to the German guideline "NVL Asthma" and international guideline GINA [1, 2], a minimum of three months of therapy with ICS is recommended at the time of initial diagnosis (maintenance therapy) before a dose reduction can be started (stepping down). Accordingly, after three months, all patients with a positive BP test or with the diagnosis of asthma, respectively, will be asked to return to the practice and examined with WBP. If asthma has been diagnosed at t1 based on BP, BP will be repeated at t2 (if the result of the bodyplethysmographic examination is inconspicuous). ICS-responsiveness is diagnosed when an airway obstruction is reversible or the tolerance to BP increases by at least one level ("doubling dose"). In addition, a potential improvement in respiratory symptoms is assessed by the Asthma Control Questionnaire (ACQ).

Approximately 2% of patients can be diagnosed with false negatives by BP tests (negative predictive value of BP determined in WBP: 98% [12]). Therefore, patients with an inconspicuous BP test are phoned after three months in order to rule out a false negative test result. Patients will be interviewed regarding symptoms and inhaler medication (structured telephone interview). An interview will take about 5 minutes. If patients report persistent respiratory symptoms although the BP test was negative, they will be offered a follow-up examination at the practice of the respective pneumologist. Depending on the findings, another BP test assessed by WBP will be performed. This will be decided by the pneumologist in each individual case.

Diagnostic decision making

A committee of experts (Antonius Schneider, member of the author board of the NVL Asthma; Rudolf A. Jörres, Senior Scientist for Respiratory Diseases, Occupational Medicine, LMU; Konrad Schultz, Medical Director of the Rehabilitation Clinic for Pneumology Bad Reichenhall) will review each diagnosis in consideration of the patient's medical history, WBP investigation, and BP. The respective pneumologists will be contacted in each inconsistent case to clarify the diagnosis. In addition, the committee of experts assesses whether the patients responded to ICS (delayed type of diagnostic study [16]).

At least one criterion must be given for an asthma diagnosis at t1:

1. increase of FEV_1 from baseline by > 12% and by > 200ml during bronchodilation testing if airway obstruction exists (NVL Asthma [1])
2. positive response of FEV_1 or Raw or sRaw during BP test

At least one of the following criteria at t2 must be fulfilled to establish ICS responsiveness at t2:

1. increase of FEV_1 from baseline by > 12% and by > 200ml (NVL asthma [1])
2. increase of tolerance during BP tests by at least one level

If criterion 1 is fulfilled, a BP test is not performed. In addition, it is not performed if the patient reports a worsening of respiratory symptoms since the initial presentation at t1.

Blinding

The FeNO measurements are performed by a doctoral candidate, a research assistant, or a lung function assistant and are documented on a structured sheet. The pneumologist who assesses the results of WBP and BP tests, is blinded to the results of the FeNO measurement. The results of the examinations and the diagnosis made by the pneumologist are documented on a separate sheet.

The committee of experts (who finally diagnoses or excludes asthma in each individual case and assesses whether the patient responded to ICS) is also blinded to the results of FeNO measurement. The committee only has access to the results of bodyplethysmographic measurements, BP tests, and anamnestic information.

Data management and monitoring

Immediately after signing the patient information, consent and data protection declaration, a pseudonymised study ID is assigned to the patient, under which the further data and study results are documented and stored. From now, all other personal data and findings will only be passed on in encrypted form, i.e. neither the name nor the initials nor the exact date of birth will appear in the encryption code. The patient identification list remains at the Institute of General Practice and Health Services Research and is only accessible to authorized study personnel. The doctoral candidate enters all data from the patient's files, the values of the FeNO measurement, and the values of the lung function tests obtained by WBP in encrypted form into the statistical program SPSS. 5% of the data will be entered twice to estimate the frequency of typing errors. Moreover, all FeNO values and all asthma diagnoses are entered twice to allow a complete correction of possible typing errors. If no more corrections are required in the database, it is closed and will be used for statistical evaluation. The data

collection process and the study procedures will be supervised by a research associate who will also perform periodic visits to the practices.

Statistics

Sample size estimation

According to previous studies in practices of pneumologists, a sample size of n=300 can be expected to include about 105 patients with a new diagnosis of asthma. The prevalence in a previous study in a large lung specialist practice was 39% [5]. To be on the safe side, we assume a slightly lower prevalence of 35% for the current multi-center study. The two primary endpoints will be tested confirmatory on two-sided 5% significance levels. A hierarchical test procedure is used to control the global type-1 error at a 5% significance level. Using exact binomial tests, the expected specificity of 95% is first tested against a reference value of 90% assumed under the null hypothesis. If the test result is positive, another confirmatory test of the expected sensitivity of 35% against a reference value of 20% will follow. These tests each achieve a power of 90% with a sample sizes of 195 patients without asthma diagnosis and 105 asthma patients [25]. The total number of patients is therefore 300.

A validation of the diagnostic algorithm (FeNO, "Allergic Rhinitis" and "Wheezing") [10] is performed by means of Wilcoxon (Mann-Whitney) rank sum tests. With the sample sizes mentioned above, this test reaches a power of 80% at a two-sided and exploratory 5% significance level to detect a diagnostic accuracy of AUC = 0.60 [26].

Statistical analysis

Patients participating in the study are characterized by descriptive statistics (mean values, standard deviations, medians, minimum, maximum; absolute and relative frequencies).

As primary and confirmatory analysis, exact binomial tests of sensitivity and specificity at t1 are performed hierarchically at the predetermined FeNO cut-off value of >50ppb, each against a reference value of 90% or 20%, respectively, and at the two-sided 5% significance level. For these measures as well as for PPV and NPV, corresponding 95% confidence intervals are calculated. Fagan nomograms will be provided for the PPV and NPV to enable the exploration of post-test probabilities depending on the population specific prevalence. The distribution of diagnoses using FeNO and the reference standard will be shown in a cross-table. The reference standard is the diagnosis of asthma made by body plethysmography and bronchoprovocation if necessary. The statistics mentioned above are calculated analogously:

- in the presence of the symptoms "Allergic rhinitis" and "Wheezing" and using a FeNO cut-off value of >33ppb at t1. In addition, the area under the curve (AUC) of the receiver operating characteristic curve (ROC) is determined with a corresponding 95%

confidence interval and tested against a reference value of 0.50 using the Wilcoxon (Mann-Whitney) rank sum test.

- for the prediction of ICS responsiveness (determined at t2) using a FeNO cut-off value of >33ppb

In accordance with the secondary hypotheses, exploratory testing of the PPV values will be performed by exact binomial tests on two-sided 5% significance levels against a reference value of 70%.

Changes in the ACQ during follow-up will be estimated by secondary analyses. For this purpose, a composite endpoint related to ICS responsiveness will be developed (at least one out of these criteria must be fulfilled):

1. increase of FEV_1 from baseline by > 12% and by > 200ml (NVL asthma [1])
2. increase of tolerance during BP tests by one level
3. improvement of 0.5 in the ACQ

3. Patient and public involvement

Patients were not involved in the design of this study.

4. Discussion

The present confirmatory diagnostic study aims to prove the diagnostic benefit of FeNO measurement regarding the diagnosis of asthma. FeNO is an attractive diagnostic tool and provides a non-invasive marker of inflammatory processes in the lung [8]. In contrast, BP as the reference standard for diagnosing asthma is time-consuming, cost-intensive, often only available in specific lung function laboratories and bears a small risk of bronchospasm [22]. Therefore, it is reasonable to discuss FeNO measurement as an alternative procedure to diagnose asthma. Beyond that, there are strong hints that it has added value to determine ICS-responsiveness. Accordingly, a health technology assessment (HTA) found that the inclusion of FeNO measurement into the diagnostic pathway might increase the diagnostic cost-effectiveness [27].

Several studies have already shown a high diagnostic accuracy of FeNO for discerning asthma in patients suspected of suffering from asthma [5, 6]. In most of these studies, values of specificity were superior to those of sensitivity, suggesting that FeNO measurement is more suitable for ruling in than for ruling out the disease [9]. However, a great weakness of the studies published so far is that the optimal FeNO cut-off values were defined post hoc. This probably led to differences when estimating the diagnostic accuracy of FeNO in different studies as well as to discrepancies regarding the optimal cut-off value for diagnosing or excluding asthma. Indeed, it is known that diagnostic algorithms, including cut-off values perform better in the dataset from which they are derived, compared to a dataset with even

similar but different individuals [28]. This phenomenon can be explained, amongst other factors, by overfitting, the absence of important predictors, unsatisfactory model derivation, and differences between patient samples [29, 30]. It is therefore essential to validate pre-defined FeNO cut-off values and diagnostic algorithms based on FeNO measurements in a prospective study, e.g. in individuals outside the derivation dataset, in order to be able to determine the adequate place of FeNO measurement in the diagnosis of asthma and in routine care [1, 9]. The present confirmatory study aims to close this gap.

Due to the confirmatory character of this study, three hypotheses are proposed before the study is conducted. Firstly, we hypothesize that a FeNO cut-off value of >50ppb is suitable to diagnose asthma (sensitivity 35%, specificity 95%). Secondly, we test the validity of the assumption that asthma can be diagnosed with a certainty (PPV) of at least 70% at a FeNO value of >33ppb, if the clinical symptoms "allergic rhinitis" and "wheezing" are present. Moreover, in line with the study by Martin et al. [14], we hypothesize that a FeNO value of >33ppb can predict an ICS responsiveness with a certainty (PPV) of at least 70%. We are aware of the discussion about using FeNO measurement better to identify responsiveness to treatment rather than to label patients with a diagnosis [13]. We aim to investigate the diagnostic usefulness regarding these aspects in a confirmatory manner. Thus, the design of the study should be suitable to verify these hypotheses.

The study will be conducted prospectively by enrolling 300 diagnostic-naïve patients from three different practices of pneumologists to increase the generalisability of the study [31]. All patients will be subjected to the reference standard to establish their true diagnosis. In this context, a major strength of the study is that the diagnosis of asthma will be made rigorously on basis of BP in WBP. It has been shown previously that interpretation of BP responsiveness with WBP, including airway resistance, is superior to the interpretation solely based on FEV_1 responsiveness [12]. After 3 months, patients with an asthma diagnosis will be examined again and the asthma diagnosis will be verified by the expert team in order to ensure the diagnosis, exclude false positive findings, and determine ICS responsiveness. In parallel, patients without an asthma diagnosis will be phoned after three months and asked if respiratory symptoms still persist. Patients with persistent symptoms will be invited for re-evaluation to exclude false negative findings. This procedure enables us to determine the prognostic value of FeNO regarding the diagnosis of asthma, and to compare the diagnostic-prognostic value of FeNO with BP. The diagnosis of each patient, as well as the evaluation of ICS responsiveness, will be made by an expert team based on anamnestic information as well as on lung function measurements, including BP tests. The expert team as well as the pneumologists of the practices are blinded to the results of FeNO measurement to avoid information bias.

A limitation of the study might be that a longer course of disease could be taken into account, e.g. with a 12-month follow-up evaluation. However, this would not allow us to use the optimal

time frame of 3 months for determining ICS responsiveness. Beyond that FeNO devices from various manufacturers should be compared since it cannot be excluded that optimal cut-off values differ between devices. We think that determination of FeNO with NIOX VERO will allow a valid estimation, because it measures FeNO at a mouth flow rate of 50 mL/s over ten seconds and a pressure of 10 cm H_2O as per guideline recommendation [32], and NIOX has been used in many diagnostic studies [9]. The present study might be able to enhance the implementation of FeNO in diagnostic guidelines. However, it will not be able to assess the impact of FeNO on diagnostic decision making in routine care and patient outcomes. This point can be only clarified in a clinical impact analysis study, which will be needed in future [28, 33, 34].

5. Ethics and dissemination

The study was approved by the Ethical Committee of the Technical University of Munich (Reference number 122/20 S). Written, informed consent to participate will be obtained from all participants. The study protocol is registered in the German Clinical Trials Register (DRKS00021125, 24 June 2020). The major results of the study will be published in peer-reviewed academic journals and disseminated through conferences.

6. Trial status

Protocol version 1.0. For recruitment the following time frame is planned: First patient in July 2020, last patient in September 2021, last patient out December 2021.

7. List of abbreviations

AUC	area under the curve
BP	bronchial provocation
FeNO	fractional exhaled nitric oxide
FEV_1	forced expiratory volume in one second
FVC	forced vital capacity
ICS	inhaled corticosteroide
NO	nitric oxide
NPV	negative predictive value
PEF	peak expiratory flow
PPV	positive predictive value
Raw	airway resistance
ROC	receiver operating characteristic
sRaw	specific airway resistance
WBP	whole body plethysmography

8. References

1. Bundesärztekammer (BÄK), Kassenärztliche Bundesvereinigung (KBV), Arbeitsgemeinschaft der Wissenschaftli-chen Medizinischen Fachgesellschaften (AWMF). Nationale VersorgungsLeitlinie Asthma – Langfassung, 4.Auflage. Konsultationsfassung. 2020 [cited: 2020-06-22].www.asthma.versorgungsleitlinien.de
2. Global Initiative for Asthma. Global Strategy for Asthma Management and Prevention, 2020. Available from: www.ginasthma.org
3. Kuenzli N, Stutz EZ, Perruchoud AP, Braendli O, Tschopp J-M, Bolognini G, et al. Peak flow variability in the SAPALDIA study and its validity in screening for asthma-related conditions. American journal of respiratory and critical care medicine. 1999;160:427-34.
4. Tilemann L, Gindner L, Meyer F, Laux G, Szecsenyi J, Schneider A. Diagnostischer Wert der Peak-Flow-Variabilität bei Verdacht auf Asthma bronchiale in der Hausarztpraxis. DMW-Deutsche Medizinische Wochenschrift. 2009;134(41):2053-8.
5. Schneider A, Schwarzbach J, Faderl B, Welker L, Karsch-Volk M, Jorres RA. FENO measurement and sputum analysis for diagnosing asthma in clinical practice. Respir Med. 2013;107(2):209-16.
6. Schneider A, Faderl B, Schwarzbach J, Welker L, Karsch-Volk M, Jorres RA. Prognostic value of bronchial provocation and FENO measurement for asthma diagnosis--results of a delayed type of diagnostic study. Respir Med. 2014;108(1):34-40.
7. Bjermer L, Alving K, Diamant Z, Magnussen H, Pavord I, Piacentini G, et al. Current evidence and future research needs for FeNO measurement in respiratory diseases. Respiratory medicine. 2014;108(6):830-41.
8. Lane C, Knight D, Burgess S, Franklin P, Horak F, Legg J, et al. Epithelial inducible nitric oxide synthase activity is the major determinant of nitric oxide concentration in exhaled breath. Thorax. 2004;59(9):757-60.
9. Karrasch S, Linde K, Rucker G, Sommer H, Karsch-Volk M, Kleijnen J, et al. Accuracy of FENO for diagnosing asthma: a systematic review. Thorax. 2017;72(2):109-16.
10. Schneider A, Linde K, Reitsma JB, Steinhauser S, Rucker G. A novel statistical model for analyzing data of a systematic review generates optimal cutoff values for fractional exhaled nitric oxide for asthma diagnosis. J Clin Epidemiol. 2017;92:69-78.
11. Schneider A, Wagenpfeil G, Jorres RA, Wagenpfeil S. Influence of the practice setting on diagnostic prediction rules using FENO measurement in combination with clinical signs and symptoms of asthma. BMJ Open. 2015;5(11):e009676.

12. Schneider A, Schwarzbach J, Faderl B, Hautmann H, Jorres RA. Whole-Body Plethysmography in Suspected Asthma: A Prospective Study of Its Added Diagnostic Value in 302 Patients. Dtsch Arztebl Int. 2015;112(24):405-11.
13. Alving K. FeNO and suspected asthma: better to identify responsiveness to treatment than to label with a diagnosis. Lancet Respir Med. 2018;6(1):3-5.
14. Martin MJ, Wilson E, Gerrard-Tarpey W, Meakin G, Hearson G, McKeever TM, et al. The utility of exhaled nitric oxide in patients with suspected asthma. Thorax. 2016;71(6):562-4.
15. Price DB, Buhl R, Chan A, Freeman D, Gardener E, Godley C, et al. Fractional exhaled nitric oxide as a predictor of response to inhaled corticosteroids in patients with non-specific respiratory symptoms and insignificant bronchodilator reversibility: a randomised controlled trial. Lancet Respir Med. 2018;6(1):29-39.
16. Knottnerus JA, Muris JW. Assessment of the accuracy of diagnostic tests: the cross-sectional study. J Clin Epidemiol. 2003;56(11):1118-28.
17. Juniper EF, O'Byrne PM, Guyatt GH, Ferrie PJ, King DR. Development and validation of a questionnaire to measure asthma control. Eur Respir J. 1999;14(4):902-7.
18. Juniper EF, Svensson K, Mork AC, Stahl E. Measurement properties and interpretation of three shortened versions of the asthma control questionnaire. Respir Med. 2005;99(5):553-8.
19. Exhaled N. ATS/ERS recommendations for standardized procedures for the online and offline measurement of exhaled lower respiratory nitric oxide and nasal nitric oxide, 2005. Am J Respir Crit Care Med. 2005;171:912-30.
20. Quanjer PH, Stanojevic S, Cole TJ, Baur X, Hall GL, Culver BH, et al. Multi-ethnic reference values for spirometry for the 3–95-yr age range: the global lung function 2012 equations. Eur Respiratory Soc; 2012.
21. Merget R, Jörres RA, Heinze E, Haufs MG, Taeger D, Brüning T. Development of a 1-concentration-4-step dosimeter protocol for methacholine testing. Respiratory medicine. 2009;103(4):607-13.
22. Crapo R. Guidelines for methacholine and exercise challenge testing-1999. This official statement of the American Thoracic Society was adopted by the ATS Board of Directors, July 1999. Am J Respir Crit Care Med. 2000;161:309-29.
23. Criee CP, Sorichter S, Smith HJ, Kardos P, Merget R, Heise D, et al. Body plethysmography--its principles and clinical use. Respir Med. 2011;105(7):959-71.
24. Perpina M, Pellicer C, de Diego A, Compte L, Macian V. Diagnostic value of the bronchial provocation test with methacholine in asthma. A Bayesian analysis approach. Chest. 1993;104(1):149-54.

25. Chow S-C, Wang H, Shao J. Sample size calculations in clinical research: CRC press; 2007.
26. Noether GE. Sample size determination for some common nonparametric tests. Journal of the American Statistical Association. 1987;82(398):645-7.
27. Harnan SE, Tappenden P, Essat M, Gomersall T, Minton J, Wong R, et al. Measurement of exhaled nitric oxide concentration in asthma: a systematic review and economic evaluation of NIOX MINO, NIOX VERO and NObreath. Health technology assessment (Winchester, England). 2015;19(82):1.
28. Moons KG, Kengne AP, Grobbee DE, Royston P, Vergouwe Y, Altman DG, et al. Risk prediction models: II. External validation, model updating, and impact assessment. Heart. 2012;98(9):691-8.
29. Steyerberg EW. Clinical prediction models: Springer; 2019.
30. Toll D, Janssen K, Vergouwe Y, Moons K. Validation, updating and impact of clinical prediction rules: a review. Journal of clinical epidemiology. 2008;61(11):1085-94.
31. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. Annals of internal medicine. 1999;130(6):515-24.
32. Dweik RA, Boggs PB, Erzurum SC, Irvin CG, Leigh MW, Lundberg JO, et al. An official ATS clinical practice guideline: interpretation of exhaled nitric oxide levels (FENO) for clinical applications. American journal of respiratory and critical care medicine. 2011;184(5):602-15.
33. Moons KG, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. Bmj. 2009;338:b606.
34. Cowley LE, Farewell DM, Maguire S, Kemp AM. Methodological standards for the development and evaluation of clinical prediction rules: a review of the literature. Diagnostic and Prognostic Research. 2019;3(1):16.

9. Declarations

Authors' contributions

C.K. prepared the final study protocol, was involved in the development of the design of the study and agreed to be accountable for all aspects of the work. A.H. developed the details of the statistical analysis plan, reviewed the manuscript and commented on drafts of the final manuscript. R.J. helped with the development of the design of the study and with manuscript preparation. K.S. and B.B. contributed to the development of the study protocol and helped with writing. A.S. developed the design of the study and was involved in the development of the statistical analysis plan as well as in manuscript preparation.

Funding statement

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors

Competing interests statement

The authors declare that they have no competing interests.

Acknowledgements

We are grateful to Drs. Hellmann, Wehgartner-Winkler, Faderl, Dankelmann, Vitiello and Schlimok in Augsburg, to Dr. Weber in Starnberg, and to Dr. Powitz in Munich for the opportunity to perform the study in their practices. We also thank the technical personnel for making the study possible through logistic and organizational support. In addition, we appreciate the willingness of all participants to perform the additional FeNO measurements.

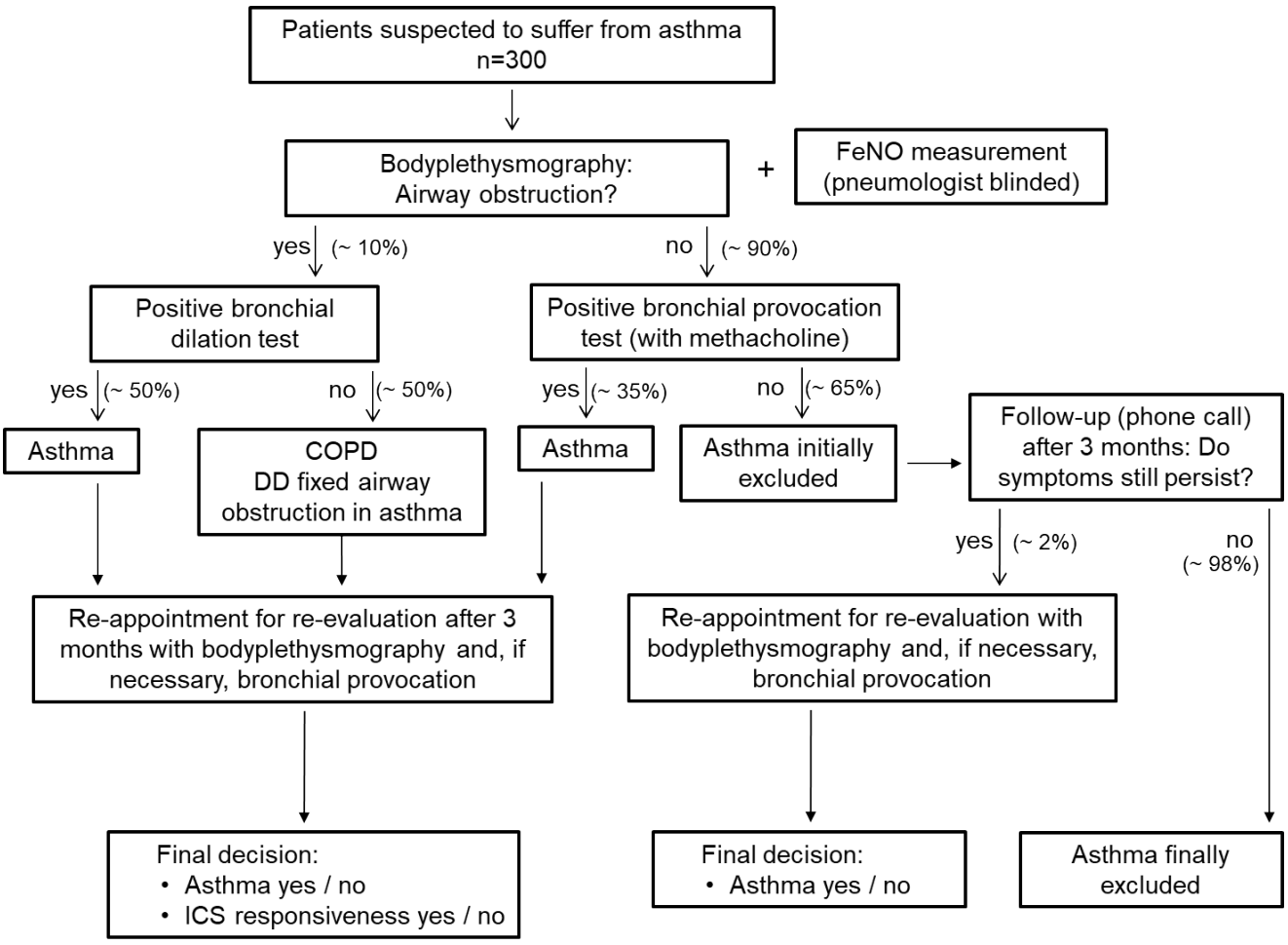


Figure 1. Overview of the diagnostic procedure during the conduct of the study.

	STUDY PERIOD			
	Enrolment	**Assessments and interventions**	**Follow-up (3 months after t1)**	
TIMEPOINT	t_0	t_1	t_2	
			Asthma diagnosis at t1	***No asthma diagnosis at t1***
ENROLMENT:				
Eligibility screen	X			
Informed consent	X			
INTERVENTIONS:				
***Index test:* FeNO**		X		
Reference test:* Bodyplethysmography and bronchial provocation test		X	X	
ASSESSMENTS:				
ACQ		X	X	
Self-reported questionnaire		X	X	
Structured interview (phone call)				X

Figure 2. Standard Protocol Items: Recommendations for interventional trials (SPIRIT) schedule. ACQ, asthma control questionnaire; FeNO, fractional exhaled nitric oxide. *bronchial provocation test is only performed if required by the pneumologist.

Section & Topic	No	Item	Reported on page #
TITLE OR ABSTRACT			
	1	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)	2
ABSTRACT			
	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	2
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the index test	3-4
	4	Study objectives and hypotheses	4
METHODS			
Study design	**5**	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)	4
Participants	**6**	Eligibility criteria	5
	7	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)	5
	8	Where and when potentially eligible participants were identified (setting, location and dates)	4-5
	9	Whether participants formed a consecutive, random or convenience series	5
Test methods	**10a**	Index test, in sufficient detail to allow replication	6
	10b	Reference standard, in sufficient detail to allow replication	6-7
	11	Rationale for choosing the reference standard (if alternatives exist)	6
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	4
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	6-7
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	8
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	8
Analysis	**14**	Methods for estimating or comparing measures of diagnostic accuracy	9
	15	How indeterminate index test or reference standard results were handled	NA
	16	How missing data on the index test and reference standard were handled	NA
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	9-10
	18	Intended sample size and how it was determined	9
RESULTS			
Participants	**19**	Flow of participants, using a diagram	Figure 1
	20	Baseline demographic and clinical characteristics of participants	NA, we are reporting the study protocol
	21a	Distribution of severity of disease in those with the target condition	NA, we are reporting the study protocol
	21b	Distribution of alternative diagnoses in those without the target condition	NA, we are reporting the study protocol
	22	Time interval and any clinical interventions between index test and reference standard	6-7
Test results	**23**	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	9-10
	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	9-10
	25	Any adverse events from performing the index test or the reference standard	NA
DISCUSSION			
	26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability	10-12
	27	Implications for practice, including the intended use and clinical role of the index test	11-12

OTHER INFORMATION			
	28	Registration number and name of registry	12
	29	Where the full study protocol can be accessed	12
	30	Sources of funding and other support; role of funders	15-16

STARD 2015

AIM

STARD stands for "Standards for Reporting Diagnostic accuracy studies". This list of items was developed to contribute to the completeness and transparency of reporting of diagnostic accuracy studies. Authors can use the list to write informative study reports. Editors and peer-reviewers can use it to evaluate whether the information has been included in manuscripts submitted for publication.

EXPLANATION

A **diagnostic accuracy study** evaluates the ability of one or more medical tests to correctly classify study participants as having a **target condition.** This can be a disease, a disease stage, response or benefit from therapy, or an event or condition in the future. A medical test can be an imaging procedure, a laboratory test, elements from history and physical examination, a combination of these, or any other method for collecting information about the current health status of a patient.

The test whose accuracy is evaluated is called **index test.** A study can evaluate the accuracy of one or more index tests. Evaluating the ability of a medical test to correctly classify patients is typically done by comparing the distribution of the index test results with those of the **reference standard**. The reference standard is the best available method for establishing the presence or absence of the target condition. An accuracy study can rely on one or more reference standards.

If test results are categorized as either positive or negative, the cross tabulation of the index test results against those of the reference standard can be used to estimate the **sensitivity** of the index test (the proportion of participants *with* the target condition who have a positive index test), and its **specificity** (the proportion *without* the target condition who have a negative index test). From this cross tabulation (sometimes referred to as the contingency or "2x2" table), several other accuracy statistics can be estimated, such as the positive and negative **predictive values** of the test. Confidence intervals around estimates of accuracy can then be calculated to quantify the statistical **precision** of the measurements.

If the index test results can take more than two values, categorization of test results as positive or negative requires a **test positivity cut-off**. When multiple such cut-offs can be defined, authors can report a receiver operating characteristic (ROC) curve which graphically represents the combination of sensitivity and specificity for each possible test positivity cut-off. The **area under the ROC curve** informs in a single numerical value about the overall diagnostic accuracy of the index test.

The **intended use** of a medical test can be diagnosis, screening, staging, monitoring, surveillance, prediction or prognosis. The **clinical role** of a test explains its position relative to existing tests in the clinical pathway. A replacement test, for example, replaces an existing test. A triage test is used before an existing test; an add-on test is used after an existing test.

Besides diagnostic accuracy, several other outcomes and statistics may be relevant in the evaluation of medical tests. Medical tests can also be used to classify patients for purposes other than diagnosis, such as staging or prognosis. The STARD list was not explicitly developed for these other outcomes, statistics, and study types, although most STARD items would still apply.

DEVELOPMENT

This STARD list was released in 2015. The 30 items were identified by an international expert group of methodologists, researchers, and editors. The guiding principle in the development of STARD was to select items that, when reported, would help readers to judge the potential for bias in the study, to appraise the applicability of the study findings and the validity of conclusions and recommendations. The list represents an update of the first version, which was published in 2003.

More information can be found on http://www.equator-network.org/reporting-guidelines/stard.

BMJ Open

Evaluation of the diagnostic accuracy of fractional exhaled nitric oxide (FeNO) in patients with suspected asthma: study protocol for a prospective diagnostic study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-045420.R1
Article Type:	Protocol
Date Submitted by the Author:	06-Dec-2020
Complete List of Authors:	Kellerer, Christina; Technical University of Munich, Institute of General Practice and Health Services Research Hapfelmeier, Alexander; Technical University of Munich, Institute of General Practice and Health Services Research Joerres, Rudolf; Ludwig-Maximilians-University Munich, Occupational and Environmental Medicine Schultz, Konrad; Bad Reichenhall Clinic, Centre for Rehabilitation, Pneumology and Orthopedics Brunn, Benjamin; Technical University of Munich, Institute of General Practice and Health Services Research Schneider, Antonius; Technical University Munich, Institute of General Practice
Primary Subject Heading:	Respiratory medicine
Secondary Subject Heading:	General practice / Family practice
Keywords:	Asthma < THORACIC MEDICINE, Chronic airways disease < THORACIC MEDICINE, GENERAL MEDICINE (see Internal Medicine)

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Evaluation of the diagnostic accuracy of fractional exhaled nitric oxide (FeNO) in patients with suspected asthma: study protocol for a prospective diagnostic study

Christina Kellerer[1]*, Alexander Hapfelmeier[1,2], Rudolf A. Joerres[3], Konrad Schultz[4], Benjamin Brunn[1], Antonius Schneider[1]

[1] Technical University of Munich, School of Medicine, Institute of General Practice and Health Services Research, Munich, Germany

[2] Institute of Medical Informatics, Statistics and Epidemiology, School of Medicine, Technical University of Munich, Munich, Germany

[3] Institute and Outpatient Clinic for Occupational, Social and Environmental Medicine, Ludwig-Maximilians-Universität München, Munich, Germany

[4] Clinic Bad Reichenhall, Center for Rehabilitation, Pneumology and Orthopedics, Bad Reichenhall

Corresponding Author:

Christina Kellerer, MSc

TUM School of Medicine

Institute of General Practice and Health Services Research

Technical University of Munich

Orleansstraße 47

81667 Munich, Germany

Mail: christina.kellerer@mri.tum.de

Phone: +49 89 6146589-17

Abstract

Introduction

The measurement of fractional exhaled nitric oxide (FeNO) is promising for diagnosing asthma and might substitute for bronchial provocation (BP) tests. To evaluate the diagnostic accuracy of FeNO within a confirmatory study, the following hypotheses will be tested: 1. A FeNO cut-off > 50ppb is suitable for diagnosing asthma (sensitivity 35%, specificity 95%). 2. If the clinical symptoms "allergic rhinitis" and "wheezing" are present, asthma can be diagnosed at FeNO > 33ppb with a positive predictive value (PPV) ≥ 70%. 3. A FeNO > 33ppb can predict responsiveness to inhaled corticosteroid (ICS) with a PPV ≥ 70%.

Methods and analysis

A prospective diagnostic study will be conducted in three practices of pneumologists in Germany. 300 patients suspected of suffering from asthma will be included. As an index test, patients perform FeNO measurement with the device NIOX VERO®. As reference a test, patients are examined with whole bodyplethysmography and BP, if necessary. After three months, patients with an asthma diagnosis will be examined again to verify the diagnosis and evaluate ICS responsiveness. Patients who did not receive an asthma diagnosis at the initial examination will be phoned after three months and asked about persistent respiratory symptoms to exclude false negative findings. As a primary target, sensitivity and specificity of FeNO > 50ppb will be determined. As a secondary target the PPV for asthma at FeNO > 33ppb, when the symptoms "allergic rhinitis" and "wheezing" are present, will be calculated. Regarding ICS responsiveness, the PPV of FeNO > 33ppb will be determined.

Ethics and dissemination

The study was approved by the Ethical Committee of the Technical University of Munich (Reference number 122/20 S). The major results will be published in peer-reviewed academic journals and disseminated through conferences.

Trial registration

German Clinical Trials Register (DRKS00021125).

Key words: FeNO, asthma, ICS responsiveness, bronchial provocation

Strengths and limitations of this study

- As this prospective confirmatory study aims to validate pre-defined FeNO cut-off values for an asthma diagnosis and ICS responsiveness it might be able to determine the appropriate place of FeNO in the diagnosis of asthma and in routine care.

- A high quality reference standard will be used in this study as the diagnosis of asthma will be made in all patients based on bronchial provocation tests assessed in whole body plethysmography and a potential asthma diagnosis will be verified after three months.
- Different devices might lead to different cut-off values. However, we are not able to compare FeNO devices from various manufacturers within this study.
- The present study is not able to assess the impact of FeNO on patient management in routine care because pneumologists will be blinded against FeNO values.

1. Introduction

Background

The diagnosis of asthma is limited by the fact that airway obstruction is often not present during investigation by spirometry or whole body plethysmography (WBP) when patients suffer from mild symptoms, thus leading to diagnostic uncertainty. For these cases, diagnostic guidelines recommend bronchial provocation (BP) tests, which can only be performed in pneumologic centres in order to diagnose or exclude asthma [1, 2]. Moreover, peak-flow variability can be assessed, but the low diagnostic value of this method has been demonstrated and it is considered as a second choice method [3, 4]. Thus, in the case of inconclusive lung function results, BP remains the reference standard for the diagnosis of asthma [1, 2].

Numerous studies have demonstrated that, in addition to BP, the measurement of fractional exhaled nitric oxide (FeNO) has a high potential for diagnosing asthma and could possibly replace BP [5, 6]. Nitric oxide (NO) is released during type-2 allergic inflammation [7] and it could be shown that patients with asthma, even in mild stages of the disease, exhale NO in higher concentrations [8]. In contrast to BP, FeNO is a non-invasive measurement that can be performed without risk to the patient in a short time.

The available studies indicate that a cut-off value of 50ppb is well suited for diagnosing asthma [9, 10]. However, such values were identified only by post-hoc analyses in the sense of multiple and exploratory testing. Accordingly, the major criticism is that the diagnostic value of the cut-off points identified and proposed so far need to be confirmed in a prospective study [1, 9].

It was shown in a secondary analysis that even lower FeNO values than 50ppb could be useful for diagnosis when considering appropriate anamnestic information. If, for example, the patient suffers from allergic rhinitis and wheezing, an asthma diagnosis can be established with a high degree of certainty when FeNO is >33 ppb [11]. However, this algorithm needs to be validated in a multicentre study. Studies also indicate that the diagnostic accuracy of FeNO measurement might be superior to BP (e.g. [9, 11]), as the latter gives correctly positive values in only about 70% of cases [12]. This might be especially true for allergic, inflammatory

alterations of respiratory tract, which might be better diagnosed via FeNO than BP [13]. In line with this, FeNO could be suitable for predicting responsiveness to inhaled corticosteroide (ICS) in asthma. The study by Martin et al. [14] showed that FeNO > 33ppb could be used to predict the response to ICS in patients with suspected asthma with a high degree of certainty. However, these values were also identified by post hoc analyses. Another study found FeNO values ≥40ppb to predict ICS responsiveness in patients with non-specific respiratory symptoms [15]. In view of these reports, it is obvious that a prospective confirmatory study is necessary to validate pre-defined cut-off values and to determine the appropriate place of FeNO in the diagnosis of asthma as well as in routine care.

Aims of the study

The present study aims to evaluate sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) to clarify the following hypotheses:

1. Primary hypothesis: The sensitivity of FeNO measurement for diagnosing asthma is 35% at the cut-off > 50ppb, and specificity is 95%.
2. Secondary hypothesis: if the clinical symptoms "allergic rhinitis" and "wheezing" are present, the PPV of FeNO > 33ppb is at least 70% (validation of the diagnostic algorithm [10]). Sensitivity, specificity and NPV will also be estimated.
3. Further secondary hypothesis: The PPV of FeNO > 33ppb for ICS responsiveness is at least 70%. Sensitivity, specificity and NPV will also be estimated.

2. Methods and analysis

Trial design

The study will be conducted as a multi-centre diagnostic study in three practices of pneumologists. Patients with suspected asthma visiting one of the three practices will be informed about the study. After having obtained the informed consent of the patient, FeNO measurement will be carried out.

Afterwards, the patient will be routinely examined with WBP and, if necessary, BP to clarify a potential asthma diagnosis. This diagnostic procedure is routinely performed in German practices of pneumologists in ambulatory care if asthma is suspected.

Three months after inclusion, patients who have been diagnosed as suffering from asthma will be invited into the practices for a follow-up examination. Patients will perform FeNO measurement and afterwards they will be again examined with WBP and BP (when appropriate). Based on the examination with WBP and BP (when appropriate), it will be verified whether the patient has responded positively to ICS (delayed type of diagnostic study [16]). Based on the recommendations of the national [1] and international [2] asthma guidelines, this time interval is reasonable for a therapy of at least three months after initial diagnosis [1, 2].

Patients who did not receive an asthma diagnosis will be phoned after three months and asked whether the respiratory symptoms still persist in order to rule out false negative findings. Patients with persistent symptoms will be invited back for re-evaluation.

Study setting

The study will be conducted in three private practices of pneumologists in Germany ("Zentrum für Pneumologie, Onkologie und Schlafmedizin am Diakonissenkrankenhaus" in Augsburg; "Lungenpraxis Starnberg" in Starnberg; "Pneumologie Elisenhof" in Munich). Further practices will be included if necessary for sufficient recruitment within the intended time frame.

Eligibility criteria

Inclusion criteria:

All adult patients suspected of suffering from asthma, who visit one of the three participating practices of pneumologists and declare their written informed consent to participate in the study will be included consecutively. Patients will be included regardless of the severity of respiratory symptoms.

Exclusion criteria:

Patients with the following criteria are excluded:

- Patients who do not agree to participate in this study
- Patients younger than 18 years (legal grounds)
- Patients who do not understand the meaning of the study due to a lack of knowledge of the German language
- Patients with already diagnosed obstructive airway disease
- Patients who smoked on the day of the examination (distortion of the FeNO results and reactivity during BP testing)
- Nitrate-rich meal (e.g. salad) before the examination (false high FeNO values)
- Patients with respiratory infection < 6 weeks before examination (distortion of the FeNO results and/or BP)

Recruitment and taking informed consent

Patients visiting one of the three participating practices of pneumologists will be contacted by a doctoral candidate (BB) from the Medical Faculty of the Technical University of Munich or a Research Associate at the Institute of General Practice and Health Services Research of the Technical University of Munich regarding possible participation in the study. They will check the inclusion and exclusion criteria and will inform the patient about the study. Finally, the attending pneumologist will provide detailed information about the study. In the case of patients who are screened for participation but who ultimately do not participate in the study (due to

disagreement or other reasons), age, gender, and the reason for non-participation will be documented anonymously in order to be able to conduct a non-responder analysis to assess a potential recruitment bias.

Interventions

Patients included in the study will be examined at first contact (time point t1) and 3 months later (time point t2). The diagnostic work-up is summarized in figure 1 and the patient timeline in figure 2.

Initial examination (t1):

During the first presentation for diagnostic work-up in one of the participating practices of pneumologists, patients are examined with a NO-measuring device (NIOX VERO®) as an index test. Afterwards patients will be examined with WBP as a reference standard; a BP test is performed additionally as part of the diagnostic routine, if required by the pneumologist. In addition, patients complete a questionnaire with structured questions about medical history and symptoms. The questionnaire also contains the "Asthma Control Questionnaire (ACQ)" [17]. The ACQ is used to determine the extent of asthma control (controlled, partially controlled, uncontrolled) and the responsiveness to ICS. ICS responsiveness is given if the ACQ score improves by at least 0.5 in the sense of a "minimal important difference" [14, 18].

Index test:

The index test is performed with the electrochemically-based NO-measuring device NIOX VERO®. This device is CE-certified, available in national and international markets, and is already widely used in practices. The FeNO measurements are performed once for each patient according to the recommendations of the ATS and ERS [19]. It is a non-invasive measurement since the patient only needs to take a deep breath through the device and exhale evenly. The FeNO measurements are performed by a doctoral candidate, or a research assistant, or a lung function assistant according to the instructions of the manufacturer. FeNO devices and measurements are provided by the Institute of General Practice and Health Services Research of the Technical University of Munich.

Reference test:

Following the FeNO measurement, an examination with WBP is routinely performed and, if required by the pneumologist, BP is performed as part of the diagnostic routine to rule-in or rule-out the diagnosis of asthma. In Germany, both of these assessments are routine tests and would also take place outside the study. Thus, there is no funding of these measurements.

WBP**:** WBP is considered as the reference standard used to diagnose obstructive airway diseases. An obstructive airway disease is indicated if FEV_1 and/or FEV_1/FVC are below their

lower limits of normal [20]. A reversible airway obstruction is diagnosed if the bronchodilation test is positive (ΔFEV_1 > 12% and 200ml). If there is no bronchial obstruction, BP is performed.

Bronchial provocation test (BP): BP is performed to determine bronchial hyperresponsiveness (BHR) to methacholine according to the 1-concentration-4-step dosimeter protocol [21]. This yields similar results as the ATS multi-concentration protocol [22] but offers advantages in clinical practice. The test is considered positive (indicating BHR) if FEV_1 decreases by at least 20% after inhalation of a maximum cumulative methacholine dose of 960µg, and/or if specific airway resistance (sRaw) increases simultaneously by at least 100% and to at least 2.0 kPa*s, and/or if airway resistance (Raw) increases simultaneously by at least 100% and to at least 0.5 kPa*s/L [22, 23].

Follow-up examination after 3 months (t2):

A single BP test as a reference standard for the diagnosis of asthma only reflects the situation at the time of examination. In some cases, patients with a positive BP test do not suffer from asthma (false positive), since the positive predictive value of BP is only about 70% [12, 24]. According to the German guideline “NVL Asthma” and international guideline GINA [1, 2], a minimum of three months of therapy with ICS is recommended at the time of initial diagnosis (maintenance therapy) before a dose reduction can be started (stepping down). Accordingly, after three months, all patients with a positive BP test or with the diagnosis of asthma, respectively, will be asked to return to the practice. During this follow-up appointment patients will receive a FeNO measurement and afterwards they will be examined with WBP. If asthma has been diagnosed at t1 based on BP, BP will be repeated at t2 (if the result of the bodyplethysmographic examination is inconspicuous). ICS-responsiveness is diagnosed when an airway obstruction is reversible or the tolerance to BP increases by at least one level ("doubling dose"). In addition, a potential improvement in respiratory symptoms is assessed by the Asthma Control Questionnaire (ACQ). Potential changes in FeNO values between t1 and t2 will be evaluated exploratory.

Approximately 2% of patients can be diagnosed with false negatives by BP tests (negative predictive value of BP determined in WBP: 98% [12]). Therefore, patients with an inconspicuous BP test are phoned after three months in order to rule out a false negative test result. Patients will be interviewed regarding symptoms and inhaler medication (structured telephone interview). An interview will take about 5 minutes. If patients report persistent respiratory symptoms although the BP test was negative, they will be offered a follow-up examination at the practice of the respective pneumologist. Depending on the findings, another BP test assessed by WBP will be performed. This will be decided by the pneumologist in each individual case.

Diagnostic decision making

A committee of experts (Antonius Schneider, member of the author board of the NVL Asthma; Rudolf A. Jörres, Senior Scientist for Respiratory Diseases, Occupational Medicine, LMU; Konrad Schultz, Medical Director of the Rehabilitation Clinic for Pneumology Bad Reichenhall) will review each diagnosis in consideration of the patient's medical history, WBP investigation, and BP. The respective pneumologists will be contacted in each inconsistent case to clarify the diagnosis. In addition, the committee of experts assesses whether the patients responded to ICS (delayed type of diagnostic study [16]).

At least one criterion must be given for an asthma diagnosis at t1:

1. increase of FEV_1 from baseline by > 12% and by > 200ml during bronchodilation testing if airway obstruction exists (NVL Asthma [1])
2. positive response of FEV_1 or Raw or sRaw during BP test

At least one of the following criteria at t2 must be fulfilled to establish ICS responsiveness at t2:

1. increase of FEV_1 from baseline by > 12% and by > 200ml (NVL asthma [1])
2. increase of tolerance during BP tests by at least one level

If criterion 1 is fulfilled, a BP test is not performed. In addition, it is not performed if the patient reports a worsening of respiratory symptoms since the initial presentation at t1.

Blinding

The FeNO measurements are performed by a doctoral candidate, a research assistant, or a lung function assistant and are documented on a structured sheet. The pneumologist who assesses the results of WBP and BP tests, is blinded to the results of the FeNO measurement. The results of the examinations and the diagnosis made by the pneumologist are documented on a separate sheet.

The committee of experts (who finally diagnoses or excludes asthma in each individual case and assesses whether the patient responded to ICS) is also blinded to the results of FeNO measurement. The committee only has access to the results of bodyplethysmographic measurements, BP tests, and anamnestic information.

Data management and monitoring

Immediately after signing the patient information, consent and data protection declaration, a pseudonymised study ID is assigned to the patient, under which the further data and study results are documented and stored. From now, all other personal data and findings will only be passed on in encrypted form, i.e. neither the name nor the initials nor the exact date of birth will appear in the encryption code. The patient identification list remains at the Institute of

General Practice and Health Services Research and is only accessible to authorized study personnel. The doctoral candidate enters all data from the patient's files, the values of the FeNO measurement, and the values of the lung function tests obtained by WBP in encrypted form into the statistical program SPSS. 5% of the data will be entered twice to estimate the frequency of typing errors. Moreover, all FeNO values and all asthma diagnoses are entered twice to allow a complete correction of possible typing errors. If no more corrections are required in the database, it is closed and will be used for statistical evaluation. The data collection process and the study procedures will be supervised by a research associate who will also perform periodic visits to the practices.

Statistics

Sample size estimation

According to previous studies in practices of pneumologists, a sample size of n=300 can be expected to include about 105 patients with a new diagnosis of asthma. The prevalence in a previous study in a large lung specialist practice was 39% [5]. To be on the safe side, we assume a slightly lower prevalence of 35% for the current multi-center study. The two primary endpoints will be tested confirmatory on two-sided 5% significance levels. A hierarchical test procedure is used to control the global type-1 error at a 5% significance level. Using exact binomial tests, the expected specificity of 95% is first tested against a reference value of 90% assumed under the null hypothesis. If the test result is positive, another confirmatory test of the expected sensitivity of 35% against a reference value of 20% will follow. These tests each achieve a power of 90% with a sample sizes of 195 patients without asthma diagnosis and 105 asthma patients [25]. The total number of patients is therefore 300.

A validation of the diagnostic algorithm (FeNO, "Allergic Rhinitis" and "Wheezing") [10] is performed by means of Wilcoxon (Mann-Whitney) rank sum tests. With the sample sizes mentioned above, this test reaches a power of 80% at a two-sided and exploratory 5% significance level to detect a diagnostic accuracy of AUC = 0.60 [26].

Statistical analysis

Patients participating in the study are characterized by descriptive statistics (mean values, standard deviations, medians, minimum, maximum; absolute and relative frequencies).

As primary and confirmatory analysis, exact binomial tests of sensitivity and specificity at t1 are performed hierarchically at the predetermined FeNO cut-off value of >50ppb, each against a reference value of 90% or 20%, respectively, and at the two-sided 5% significance level. For these measures as well as for PPV and NPV, corresponding 95% confidence intervals are calculated. Fagan nomograms will be provided for the PPV and NPV to enable the exploration of post-test probabilities depending on the population specific prevalence. The distribution of

diagnoses using FeNO and the reference standard will be shown in a cross-table. The reference standard is the diagnosis of asthma made by body plethysmography and bronchoprovocation if necessary. The statistics mentioned above are calculated analogously:

- in the presence of the symptoms "Allergic rhinitis" and "Wheezing" and using a FeNO cut-off value of >33ppb at t1. In addition, the area under the curve (AUC) of the receiver operating characteristic curve (ROC) is determined with a corresponding 95% confidence interval and tested against a reference value of 0.50 using the Wilcoxon (Mann-Whitney) rank sum test.
- for the prediction of ICS responsiveness (determined at t2) using a FeNO cut-off value of >33ppb

In accordance with the secondary hypotheses, exploratory testing of the PPV values will be performed by exact binomial tests on two-sided 5% significance levels against a reference value of 70%.

Changes in the ACQ during follow-up will be estimated by secondary analyses. For this purpose, a composite endpoint related to ICS responsiveness will be developed (at least one out of these criteria must be fulfilled):

1. increase of FEV_1 from baseline by > 12% and by > 200ml (NVL asthma [1])
2. increase of tolerance during BP tests by one level
3. improvement of 0.5 in the ACQ

Moreover, regarding ICS responsiveness potential changes in FeNO values between the first appointment (t1) and the follow-up appointment (t2) will be evaluated exploratory in secondary analyses. Additionally, subgroup analyses related to different cut-off values of bronchial provocation will be performed. Furthermore, the influence of anthropometric parameters on FeNO values will be analysed in secondary analyses.

3. Patient and public involvement

Patients were not involved in the design of this study.

4. Discussion

The present confirmatory diagnostic study aims to prove the diagnostic benefit of FeNO measurement regarding the diagnosis of asthma. FeNO is an attractive diagnostic tool and provides a non-invasive marker of inflammatory processes in the lung [8]. In contrast, BP as the reference standard for diagnosing asthma is time-consuming, cost-intensive, often only available in specific lung function laboratories and bears a small risk of bronchospasm [22]. Therefore, it is reasonable to discuss FeNO measurement as an alternative procedure to diagnose asthma. Beyond that, there are strong hints that it has added value to determine ICS-

responsiveness. Accordingly, a health technology assessment (HTA) found that the inclusion of FeNO measurement into the diagnostic pathway might increase the diagnostic cost-effectiveness [27].

Several studies have already shown a high diagnostic accuracy of FeNO for discerning asthma in patients suspected of suffering from asthma [5, 6]. In most of these studies, values of specificity were superior to those of sensitivity, suggesting that FeNO measurement is more suitable for ruling in than for ruling out the disease [9]. However, a great weakness of the studies published so far is that the optimal FeNO cut-off values were defined post hoc. This probably led to differences when estimating the diagnostic accuracy of FeNO in different studies as well as to discrepancies regarding the optimal cut-off value for diagnosing or excluding asthma. Indeed, it is known that diagnostic algorithms, including cut-off values perform better in the dataset from which they are derived, compared to a dataset with even similar but different individuals [28]. This phenomenon can be explained, amongst other factors, by overfitting, the absence of important predictors, unsatisfactory model derivation, and differences between patient samples [29, 30]. It is therefore essential to validate pre-defined FeNO cut-off values and diagnostic algorithms based on FeNO measurements in a prospective study, e.g. in individuals outside the derivation dataset, in order to be able to determine the adequate place of FeNO measurement in the diagnosis of asthma and in routine care [1, 9]. The present confirmatory study aims to close this gap.

Due to the confirmatory character of this study, three hypotheses are proposed before the study is conducted. Firstly, we hypothesize that a FeNO cut-off value of >50ppb is suitable to diagnose asthma (sensitivity 35%, specificity 95%). Secondly, we test the validity of the assumption that asthma can be diagnosed with a certainty (PPV) of at least 70% at a FeNO value of >33ppb, if the clinical symptoms "allergic rhinitis" and "wheezing" are present. Moreover, in line with the study by Martin et al. [14], we hypothesize that a FeNO value of >33ppb can predict an ICS responsiveness with a certainty (PPV) of at least 70%. We are aware of the discussion about using FeNO measurement better to identify responsiveness to treatment rather than to label patients with a diagnosis [13]. We aim to investigate the diagnostic usefulness regarding these aspects in a confirmatory manner. Thus, the design of the study should be suitable to verify these hypotheses.

The study will be conducted prospectively by enrolling 300 diagnostic-naïve patients from three different practices of pneumologists to increase the generalisability of the study [31]. All patients will be subjected to the reference standard to establish their true diagnosis. In this context, a major strength of the study is that the diagnosis of asthma will be made rigorously on basis of BP in WBP. It has been shown previously that interpretation of BP responsiveness with WBP, including airway resistance, is superior to the interpretation solely based on FEV_1 responsiveness [12]. After 3 months, patients with an asthma diagnosis will be examined again

and the diagnosis of asthma will be verified by the expert team in order to ensure the diagnosis, exclude false positive findings, and determine ICS responsiveness. In parallel, patients without an asthma diagnosis will be phoned after three months and asked if respiratory symptoms still persist. Patients with persistent symptoms will be invited for re-evaluation to exclude false negative findings. This procedure enables us to determine the prognostic value of FeNO regarding the diagnosis of asthma, and to compare the diagnostic-prognostic value of FeNO with BP. The diagnosis of each patient, as well as the evaluation of ICS responsiveness, will be made by an expert team based on anamnestic information as well as on lung function measurements, including BP tests. The expert team as well as the pneumologists of the practices are blinded to the results of FeNO measurement to avoid information bias.

A limitation of the study might be that a longer course of disease could be taken into account, e.g. with a 12-month follow-up evaluation. However, this would not allow us to use the optimal time frame of 3 months for determining ICS responsiveness. Moreover, another limitation of the study might be the fact that the presence of allergic rhinitis is reported by the patient without objective validation. However, this represents the typical state of knowledge in clinical practice as it is uncommon to verify the presence of allergic rhinitis with nasal provocation in pneumological practices. Moreover, it has to be mentioned that we could not include special measures to control for adherence regarding ICS inhalation and consequently this aspect cannot be controlled in this study. Beyond that FeNO devices from various manufacturers should be compared since it cannot be excluded that optimal cut-off values differ between devices. We think that determination of FeNO with NIOX VERO will allow a valid estimation, because it measures FeNO at a mouth flow rate of 50 mL/s over ten seconds and a pressure of 10 cm H_2O as per guideline recommendation [32], and NIOX has been used in many diagnostic studies [9]. The present study might be able to enhance the implementation of FeNO in diagnostic guidelines. However, it will not be able to assess the impact of FeNO on diagnostic decision making in routine care and patient outcomes. This point can be only clarified in a clinical impact analysis study, which will be needed in future [28, 33, 34].

5. Ethics and dissemination

The study was approved by the Ethical Committee of the Technical University of Munich (Reference number 122/20 S). Written, informed consent to participate will be obtained from all participants. The study protocol is registered in the German Clinical Trials Register (DRKS00021125, 24 June 2020). The major results of the study will be published in peer-reviewed academic journals and disseminated through conferences.

6. Trial status

Protocol version 1.0. For recruitment the following time frame is planned: First patient in July 2020, last patient in September 2021, last patient out December 2021.

7. List of abbreviations

AUC	area under the curve
BP	bronchial provocation
FeNO	fractional exhaled nitric oxide
FEV_1	forced expiratory volume in one second
FVC	forced vital capacity
ICS	inhaled corticosteroide
NO	nitric oxide
NPV	negative predictive value
PEF	peak expiratory flow
PPV	positive predictive value
Raw	airway resistance
ROC	receiver operating characteristic
sRaw	specific airway resistance
WBP	whole body plethysmography

8. Declarations

Authors' contributions

C.K. prepared the final study protocol, was involved in the development of the design of the study and agreed to be accountable for all aspects of the work. A.H. developed the details of the statistical analysis plan, reviewed the manuscript and commented on drafts of the final manuscript. R.J. helped with the development of the design of the study and with manuscript preparation. K.S. and B.B. contributed to the development of the study protocol and helped with writing. A.S. developed the design of the study and was involved in the development of the statistical analysis plan as well as in manuscript preparation.

Funding statement

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests statement

The authors declare that they have no competing interests.

Acknowledgements

We are grateful to Drs. Hellmann, Wehgartner-Winkler, Faderl, Dankelmann, Vitiello and Schlimok in Augsburg, to Dr. Weber in Starnberg, and to Dr. Powitz in Munich for the opportunity to perform the study in their practices. We also thank the technical personnel for making the study possible through logistic and organizational support. In addition, we appreciate the willingness of all participants to perform the additional FeNO measurements.

9. References

1. Bundesärztekammer (BÄK), Kassenärztliche Bundesvereinigung (KBV), Arbeitsgemeinschaft der Wissenschaftli-chen Medizinischen Fachgesellschaften (AWMF). Nationale VersorgungsLeitlinie Asthma – Langfassung, 4.Auflage. Konsultationsfassung. 2020 [cited: 2020-06-22].www.asthma.versorgungsleitlinien.de
2. Global Initiative for Asthma. Global Strategy for Asthma Management and Prevention, 2020. Available from: www.ginasthma.org
3. Kuenzli N, Stutz EZ, Perruchoud AP, Braendli O, Tschopp J-M, Bolognini G, et al. Peak flow variability in the SAPALDIA study and its validity in screening for asthma-related conditions. American journal of respiratory and critical care medicine. 1999;160:427-34.
4. Tilemann L, Gindner L, Meyer F, Laux G, Szecsenyi J, Schneider A. Diagnostischer Wert der Peak-Flow-Variabilität bei Verdacht auf Asthma bronchiale in der Hausarztpraxis. DMW-Deutsche Medizinische Wochenschrift. 2009;134(41):2053-8.
5. Schneider A, Schwarzbach J, Faderl B, Welker L, Karsch-Volk M, Jorres RA. FENO measurement and sputum analysis for diagnosing asthma in clinical practice. Respir Med. 2013;107(2):209-16.
6. Schneider A, Faderl B, Schwarzbach J, Welker L, Karsch-Volk M, Jorres RA. Prognostic value of bronchial provocation and FENO measurement for asthma diagnosis--results of a delayed type of diagnostic study. Respir Med. 2014;108(1):34-40.
7. Bjermer L, Alving K, Diamant Z, Magnussen H, Pavord I, Piacentini G, et al. Current evidence and future research needs for FeNO measurement in respiratory diseases. Respiratory medicine. 2014;108(6):830-41.
8. Lane C, Knight D, Burgess S, Franklin P, Horak F, Legg J, et al. Epithelial inducible nitric oxide synthase activity is the major determinant of nitric oxide concentration in exhaled breath. Thorax. 2004;59(9):757-60.

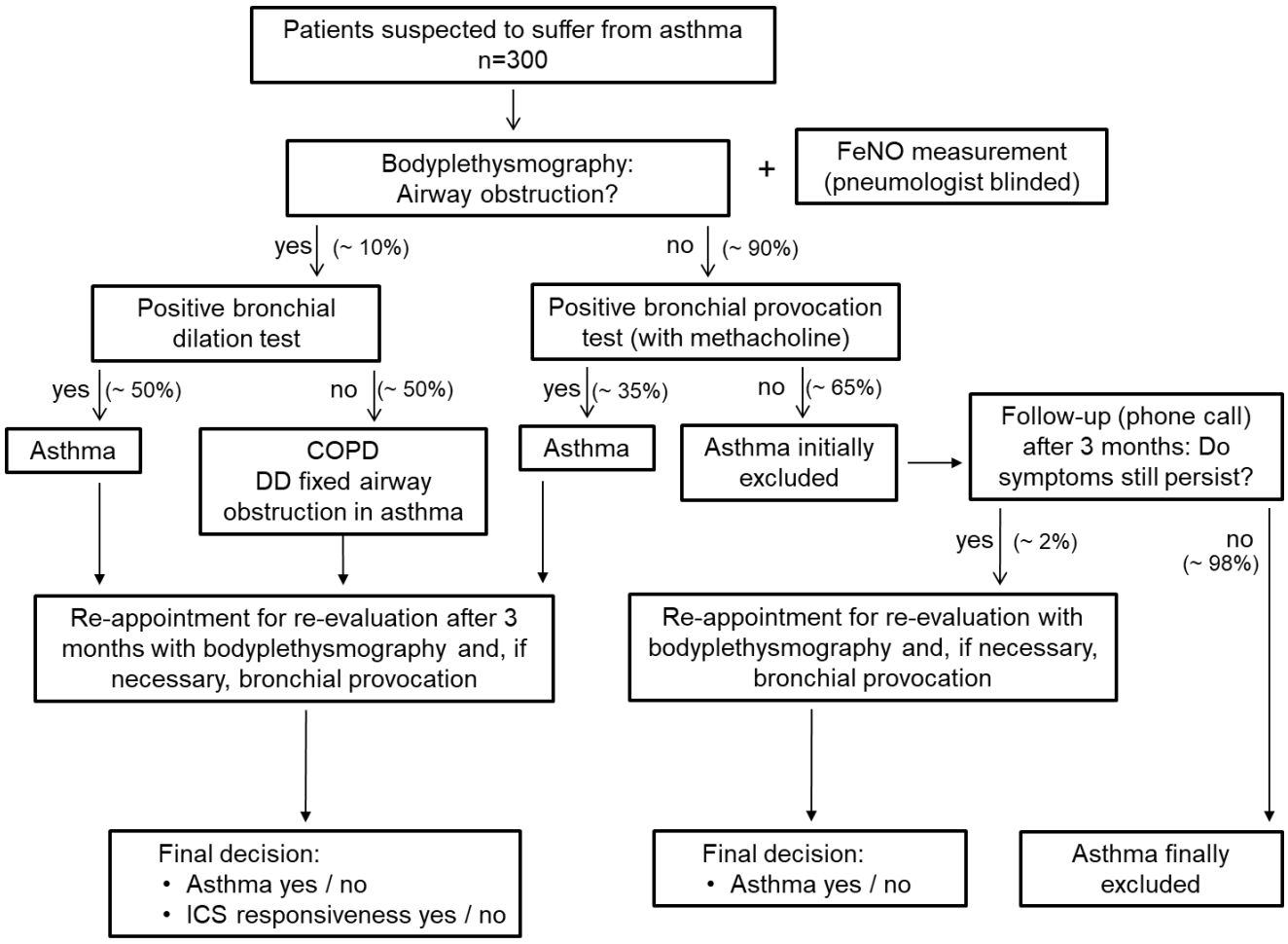
9. Karrasch S, Linde K, Rucker G, Sommer H, Karsch-Volk M, Kleijnen J, et al. Accuracy of FENO for diagnosing asthma: a systematic review. Thorax. 2017;72(2):109-16.
10. Schneider A, Linde K, Reitsma JB, Steinhauser S, Rucker G. A novel statistical model for analyzing data of a systematic review generates optimal cutoff values for fractional exhaled nitric oxide for asthma diagnosis. J Clin Epidemiol. 2017;92:69-78.
11. Schneider A, Wagenpfeil G, Jorres RA, Wagenpfeil S. Influence of the practice setting on diagnostic prediction rules using FENO measurement in combination with clinical signs and symptoms of asthma. BMJ Open. 2015;5(11):e009676.
12. Schneider A, Schwarzbach J, Faderl B, Hautmann H, Jorres RA. Whole-Body Plethysmography in Suspected Asthma: A Prospective Study of Its Added Diagnostic Value in 302 Patients. Dtsch Arztebl Int. 2015;112(24):405-11.
13. Alving K. FeNO and suspected asthma: better to identify responsiveness to treatment than to label with a diagnosis. Lancet Respir Med. 2018;6(1):3-5.
14. Martin MJ, Wilson E, Gerrard-Tarpey W, Meakin G, Hearson G, McKeever TM, et al. The utility of exhaled nitric oxide in patients with suspected asthma. Thorax. 2016;71(6):562-4.
15. Price DB, Buhl R, Chan A, Freeman D, Gardener E, Godley C, et al. Fractional exhaled nitric oxide as a predictor of response to inhaled corticosteroids in patients with non-specific respiratory symptoms and insignificant bronchodilator reversibility: a randomised controlled trial. Lancet Respir Med. 2018;6(1):29-39.
16. Knottnerus JA, Muris JW. Assessment of the accuracy of diagnostic tests: the cross-sectional study. J Clin Epidemiol. 2003;56(11):1118-28.
17. Juniper EF, O'Byrne PM, Guyatt GH, Ferrie PJ, King DR. Development and validation of a questionnaire to measure asthma control. Eur Respir J. 1999;14(4):902-7.
18. Juniper EF, Svensson K, Mork AC, Stahl E. Measurement properties and interpretation of three shortened versions of the asthma control questionnaire. Respir Med. 2005;99(5):553-8.
19. Exhaled N. ATS/ERS recommendations for standardized procedures for the online and offline measurement of exhaled lower respiratory nitric oxide and nasal nitric oxide, 2005. Am J Respir Crit Care Med. 2005;171:912-30.
20. Quanjer PH, Stanojevic S, Cole TJ, Baur X, Hall GL, Culver BH, et al. Multi-ethnic reference values for spirometry for the 3–95-yr age range: the global lung function 2012 equations. Eur Respiratory Soc; 2012.
21. Merget R, Jörres RA, Heinze E, Haufs MG, Taeger D, Brüning T. Development of a 1-concentration-4-step dosimeter protocol for methacholine testing. Respiratory medicine. 2009;103(4):607-13.

22. Crapo R. Guidelines for methacholine and exercise challenge testing-1999. This official statement of the American Thoracic Society was adopted by the ATS Board of Directors, July 1999. Am J Respir Crit Care Med. 2000;161:309-29.
23. Criee CP, Sorichter S, Smith HJ, Kardos P, Merget R, Heise D, et al. Body plethysmography--its principles and clinical use. Respir Med. 2011;105(7):959-71.
24. Perpina M, Pellicer C, de Diego A, Compte L, Macian V. Diagnostic value of the bronchial provocation test with methacholine in asthma. A Bayesian analysis approach. Chest. 1993;104(1):149-54.
25. Chow S-C, Wang H, Shao J. Sample size calculations in clinical research: CRC press; 2007.
26. Noether GE. Sample size determination for some common nonparametric tests. Journal of the American Statistical Association. 1987;82(398):645-7.
27. Harnan SE, Tappenden P, Essat M, Gomersall T, Minton J, Wong R, et al. Measurement of exhaled nitric oxide concentration in asthma: a systematic review and economic evaluation of NIOX MINO, NIOX VERO and NObreath. Health technology assessment (Winchester, England). 2015;19(82):1.
28. Moons KG, Kengne AP, Grobbee DE, Royston P, Vergouwe Y, Altman DG, et al. Risk prediction models: II. External validation, model updating, and impact assessment. Heart. 2012;98(9):691-8.
29. Steyerberg EW. Clinical prediction models: Springer; 2019.
30. Toll D, Janssen K, Vergouwe Y, Moons K. Validation, updating and impact of clinical prediction rules: a review. Journal of clinical epidemiology. 2008;61(11):1085-94.
31. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. Annals of internal medicine. 1999;130(6):515-24.
32. Dweik RA, Boggs PB, Erzurum SC, Irvin CG, Leigh MW, Lundberg JO, et al. An official ATS clinical practice guideline: interpretation of exhaled nitric oxide levels (FENO) for clinical applications. American journal of respiratory and critical care medicine. 2011;184(5):602-15.
33. Moons KG, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. Bmj. 2009;338:b606.
34. Cowley LE, Farewell DM, Maguire S, Kemp AM. Methodological standards for the development and evaluation of clinical prediction rules: a review of the literature. Diagnostic and Prognostic Research. 2019;3(1):16.

Legends to figures

Figure 1. Overview of the diagnostic procedure during the conduct of the study.

Figure 2. Standard Protocol Items: Recommendations for interventional trials (SPIRIT) schedule. ACQ, asthma control questionnaire; FeNO, fractional exhaled nitric oxide. *bronchial provocation test is only performed if required by the pneumologist.



	STUDY PERIOD			
	Enrolment	**Assessments and interventions**	**Follow-up (3 months after t1)**	
TIMEPOINT	t_0	t_1	t_2	
			Asthma diagnosis at t1	***No asthma diagnosis at t1***
ENROLMENT:				
Eligibility screen	X			
Informed consent	X			
INTERVENTIONS:				
***Index test:* FeNO**		X		
Reference test:* Bodyplethysmography and bronchial provocation test		X	X	
ASSESSMENTS:				
ACQ		X	X	
Self-reported questionnaire		X	X	
Structured interview (phone call)				X

Section & Topic	No	Item	Reported on page #
TITLE OR ABSTRACT			
	1	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)	2
ABSTRACT			
	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	2
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the index test	3-4
	4	Study objectives and hypotheses	4
METHODS			
Study design	**5**	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)	4
Participants	**6**	Eligibility criteria	5
	7	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)	5
	8	Where and when potentially eligible participants were identified (setting, location and dates)	4-5
	9	Whether participants formed a consecutive, random or convenience series	5
Test methods	**10a**	Index test, in sufficient detail to allow replication	6
	10b	Reference standard, in sufficient detail to allow replication	6-7
	11	Rationale for choosing the reference standard (if alternatives exist)	6
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	4
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	6-7
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	8
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	8
Analysis	**14**	Methods for estimating or comparing measures of diagnostic accuracy	9
	15	How indeterminate index test or reference standard results were handled	NA
	16	How missing data on the index test and reference standard were handled	NA
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	9-10
	18	Intended sample size and how it was determined	9
RESULTS			
Participants	**19**	Flow of participants, using a diagram	Figure 1
	20	Baseline demographic and clinical characteristics of participants	NA, we are reporting the study protocol
	21a	Distribution of severity of disease in those with the target condition	NA, we are reporting the study protocol
	21b	Distribution of alternative diagnoses in those without the target condition	NA, we are reporting the study protocol
	22	Time interval and any clinical interventions between index test and reference standard	6-7
Test results	**23**	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	9-10
	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	9-10
	25	Any adverse events from performing the index test or the reference standard	NA
DISCUSSION			
	26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability	10-12
	27	Implications for practice, including the intended use and clinical role of the index test	11-12

OTHER INFORMATION			
	28	Registration number and name of registry	12
	29	Where the full study protocol can be accessed	12
	30	Sources of funding and other support; role of funders	15-16

STARD 2015

AIM

STARD stands for "Standards for Reporting Diagnostic accuracy studies". This list of items was developed to contribute to the completeness and transparency of reporting of diagnostic accuracy studies. Authors can use the list to write informative study reports. Editors and peer-reviewers can use it to evaluate whether the information has been included in manuscripts submitted for publication.

EXPLANATION

A **diagnostic accuracy study** evaluates the ability of one or more medical tests to correctly classify study participants as having a **target condition.** This can be a disease, a disease stage, response or benefit from therapy, or an event or condition in the future. A medical test can be an imaging procedure, a laboratory test, elements from history and physical examination, a combination of these, or any other method for collecting information about the current health status of a patient.

The test whose accuracy is evaluated is called **index test.** A study can evaluate the accuracy of one or more index tests. Evaluating the ability of a medical test to correctly classify patients is typically done by comparing the distribution of the index test results with those of the **reference standard**. The reference standard is the best available method for establishing the presence or absence of the target condition. An accuracy study can rely on one or more reference standards.

If test results are categorized as either positive or negative, the cross tabulation of the index test results against those of the reference standard can be used to estimate the **sensitivity** of the index test (the proportion of participants *with* the target condition who have a positive index test), and its **specificity** (the proportion *without* the target condition who have a negative index test). From this cross tabulation (sometimes referred to as the contingency or "2x2" table), several other accuracy statistics can be estimated, such as the positive and negative **predictive values** of the test. Confidence intervals around estimates of accuracy can then be calculated to quantify the statistical **precision** of the measurements.

If the index test results can take more than two values, categorization of test results as positive or negative requires a **test positivity cut-off**. When multiple such cut-offs can be defined, authors can report a receiver operating characteristic (ROC) curve which graphically represents the combination of sensitivity and specificity for each possible test positivity cut-off. The **area under the ROC curve** informs in a single numerical value about the overall diagnostic accuracy of the index test.

The **intended use** of a medical test can be diagnosis, screening, staging, monitoring, surveillance, prediction or prognosis. The **clinical role** of a test explains its position relative to existing tests in the clinical pathway. A replacement test, for example, replaces an existing test. A triage test is used before an existing test; an add-on test is used after an existing test.

Besides diagnostic accuracy, several other outcomes and statistics may be relevant in the evaluation of medical tests. Medical tests can also be used to classify patients for purposes other than diagnosis, such as staging or prognosis. The STARD list was not explicitly developed for these other outcomes, statistics, and study types, although most STARD items would still apply.

DEVELOPMENT

This STARD list was released in 2015. The 30 items were identified by an international expert group of methodologists, researchers, and editors. The guiding principle in the development of STARD was to select items that, when reported, would help readers to judge the potential for bias in the study, to appraise the applicability of the study findings and the validity of conclusions and recommendations. The list represents an update of the first version, which was published in 2003.

More information can be found on http://www.equator-network.org/reporting-guidelines/stard.

BMJ Open

Evaluation of the diagnostic accuracy of fractional exhaled nitric oxide (FeNO) in patients with suspected asthma: study protocol for a prospective diagnostic study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-045420.R2
Article Type:	Protocol
Date Submitted by the Author:	28-Jan-2021
Complete List of Authors:	Kellerer, Christina; Technical University of Munich, Institute of General Practice and Health Services Research Hapfelmeier, Alexander; Technical University of Munich, Institute of General Practice and Health Services Research Joerres, Rudolf; Ludwig-Maximilians-University Munich, Occupational and Environmental Medicine Schultz, Konrad; Bad Reichenhall Clinic, Centre for Rehabilitation, Pneumology and Orthopedics Brunn, Benjamin; Technical University of Munich, Institute of General Practice and Health Services Research Schneider, Antonius; Technical University Munich, Institute of General Practice
Primary Subject Heading:	Respiratory medicine
Secondary Subject Heading:	General practice / Family practice
Keywords:	Asthma < THORACIC MEDICINE, Chronic airways disease < THORACIC MEDICINE, GENERAL MEDICINE (see Internal Medicine)

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Evaluation of the diagnostic accuracy of fractional exhaled nitric oxide (FeNO) in patients with suspected asthma: study protocol for a prospective diagnostic study

Christina Kellerer[1]*, Alexander Hapfelmeier[1,2], Rudolf A. Jörres[3], Konrad Schultz[4], Benjamin Brunn[1], Antonius Schneider[1]

[1] Technical University of Munich, School of Medicine, Institute of General Practice and Health Services Research, Munich, Germany

[2] Institute of Medical Informatics, Statistics and Epidemiology, School of Medicine, Technical University of Munich, Munich, Germany

[3] Institute and Outpatient Clinic for Occupational, Social and Environmental Medicine, Ludwig-Maximilians-Universität München, Munich, Germany

[4] Clinic Bad Reichenhall, Center for Rehabilitation, Pneumology and Orthopedics, Bad Reichenhall

Corresponding Author:

Christina Kellerer, MSc

TUM School of Medicine

Institute of General Practice and Health Services Research

Technical University of Munich

Orleansstraße 47

81667 Munich, Germany

Mail: christina.kellerer@mri.tum.de

Phone: +49 89 6146589-17

Abstract

Introduction

The measurement of fractional exhaled nitric oxide (FeNO) is promising for diagnosing asthma and might substitute for bronchial provocation (BP) tests. To evaluate the diagnostic accuracy of FeNO within a confirmatory study, the following hypotheses will be tested: 1. A FeNO cut-off > 50ppb is suitable for diagnosing asthma (sensitivity 35%, specificity 95%). 2. If the clinical symptoms "allergic rhinitis" and "wheezing" are present, asthma can be diagnosed at FeNO > 33ppb with a positive predictive value (PPV) ≥ 70%. 3. A FeNO > 33ppb can predict responsiveness to inhaled corticosteroid (ICS) with a PPV ≥ 70%.

Methods and analysis

A prospective diagnostic study will be conducted in three practices of pneumologists in Germany. 300 patients suspected of suffering from asthma will be included. As an index test, patients perform FeNO measurement with the device NIOX VERO®. As reference a test, patients are examined with whole bodyplethysmography and BP, if necessary. After three months, patients with an asthma diagnosis will be examined again to verify the diagnosis and evaluate ICS responsiveness. Patients who did not receive an asthma diagnosis at the initial examination will be phoned after three months and asked about persistent respiratory symptoms to exclude false negative findings. As a primary target, sensitivity and specificity of FeNO > 50ppb will be determined. As a secondary target the PPV for asthma at FeNO > 33ppb, when the symptoms "allergic rhinitis" and "wheezing" are present, will be calculated. Regarding ICS responsiveness, the PPV of FeNO > 33ppb will be determined.

Ethics and dissemination

The study was approved by the Ethical Committee of the Technical University of Munich (Reference number 122/20 S). The major results will be published in peer-reviewed academic journals and disseminated through conferences.

Trial registration

German Clinical Trials Register (DRKS00021125).

Key words: FeNO, asthma, ICS responsiveness, bronchial provocation

Strengths and limitations of this study

- As this prospective confirmatory study aims to validate pre-defined FeNO cut-off values for an asthma diagnosis and ICS responsiveness it might be able to determine the appropriate place of FeNO in the diagnosis of asthma and in routine care.

- A high quality reference standard will be used in this study as the diagnosis of asthma will be made in all patients based on bronchial provocation tests assessed in whole body plethysmography and a potential asthma diagnosis will be verified after three months.
- Different devices might lead to different cut-off values. However, we are not able to compare FeNO devices from various manufacturers within this study.
- The present study is not able to assess the impact of FeNO on patient management in routine care because pneumologists will be blinded against FeNO values.

1. Introduction

Background

The diagnosis of asthma is limited by the fact that airway obstruction is often not present during investigation by spirometry or whole body plethysmography (WBP) when patients suffer from mild symptoms, thus leading to diagnostic uncertainty. For these cases, diagnostic guidelines recommend bronchial provocation (BP) tests, which can only be performed in pneumologic centres in order to diagnose or exclude asthma [1, 2]. Moreover, peak-flow variability can be assessed, but the low diagnostic value of this method has been demonstrated and it is considered as a second choice method [3, 4]. Thus, in the case of inconclusive lung function results, BP remains the reference standard for the diagnosis of asthma [1, 2].

Numerous studies have demonstrated that, in addition to BP, the measurement of fractional exhaled nitric oxide (FeNO) has a high potential for diagnosing asthma and could possibly replace BP [5, 6]. Nitric oxide (NO) is released during type-2 allergic inflammation [7] and it could be shown that patients with asthma, even in mild stages of the disease, exhale NO in higher concentrations [8]. In contrast to BP, FeNO is a non-invasive measurement that can be performed without risk to the patient in a short time.

The available studies indicate that a cut-off value of 50ppb is well suited for diagnosing asthma [9, 10]. However, such values were identified only by post-hoc analyses in the sense of multiple and exploratory testing. Accordingly, the major criticism is that the diagnostic value of the cut-off points identified and proposed so far need to be confirmed in a prospective study [1, 9].

It was shown in a secondary analysis that even lower FeNO values than 50ppb could be useful for diagnosis when considering appropriate anamnestic information. If, for example, the patient suffers from allergic rhinitis and wheezing, an asthma diagnosis can be established with a high degree of certainty when FeNO is >33 ppb [11]. However, this algorithm needs to be validated in a multicentre study. Studies also indicate that the diagnostic accuracy of FeNO measurement might be superior to BP (e.g. [9, 11]), as the latter gives correctly positive values in only about 70% of cases [12]. This might be especially true for allergic, inflammatory

alterations of respiratory tract, which might be better diagnosed via FeNO than BP [13]. In line with this, FeNO could be suitable for predicting responsiveness to inhaled corticosteroide (ICS) in asthma. The study by Martin et al. [14] showed that FeNO > 33ppb could be used to predict the response to ICS in patients with suspected asthma with a high degree of certainty. However, these values were also identified by post hoc analyses. Another study found FeNO values ≥40ppb to predict ICS responsiveness in patients with non-specific respiratory symptoms [15]. In view of these reports, it is obvious that a prospective confirmatory study is necessary to validate pre-defined cut-off values and to determine the appropriate place of FeNO in the diagnosis of asthma as well as in routine care.

Aims of the study

The present study aims to evaluate sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) to clarify the following hypotheses:

1. Primary hypothesis: The sensitivity of FeNO measurement for diagnosing asthma is 35% at the cut-off > 50ppb, and specificity is 95%.
2. Secondary hypothesis: if the clinical symptoms "allergic rhinitis" and "wheezing" are present, the PPV of FeNO > 33ppb is at least 70% (validation of the diagnostic algorithm [10]). Sensitivity, specificity and NPV will also be estimated.
3. Further secondary hypothesis: The PPV of FeNO > 33ppb for ICS responsiveness is at least 70%. Sensitivity, specificity and NPV will also be estimated.

2. Methods and analysis

Trial design

The study will be conducted as a multi-centre diagnostic study in three practices of pneumologists. Patients with suspected asthma visiting one of the three practices will be informed about the study. After having obtained the informed consent of the patient, FeNO measurement will be carried out.

Afterwards, the patient will be routinely examined with WBP and, if necessary, BP to clarify a potential asthma diagnosis. This diagnostic procedure is routinely performed in German practices of pneumologists in ambulatory care if asthma is suspected.

Three months after inclusion, patients who have been diagnosed as suffering from asthma will be invited into the practices for a follow-up examination. Patients will perform FeNO measurement and afterwards they will be again examined with WBP and BP (when appropriate). Based on the examination with WBP and BP (when appropriate), it will be verified whether the patient has responded positively to ICS (delayed type of diagnostic study [16]). Based on the recommendations of the national [1] and international [2] asthma guidelines, this time interval is reasonable for a therapy of at least three months after initial diagnosis [1, 2].

Patients who did not receive an asthma diagnosis will be phoned after three months and asked whether the respiratory symptoms still persist in order to rule out false negative findings. Patients with persistent symptoms will be invited back for re-evaluation.

Study setting

The study will be conducted in three private practices of pneumologists in Germany ("Zentrum für Pneumologie, Onkologie und Schlafmedizin am Diakonissenkrankenhaus" in Augsburg; "Lungenpraxis Starnberg" in Starnberg; "Pneumologie Elisenhof" in Munich). Further practices will be included if necessary for sufficient recruitment within the intended time frame.

Eligibility criteria

Inclusion criteria:

All adult patients suspected of suffering from asthma, who visit one of the three participating practices of pneumologists and declare their written informed consent to participate in the study will be included consecutively. Patients will be included regardless of the severity of respiratory symptoms.

Exclusion criteria:

Patients with the following criteria are excluded:

- Patients who do not agree to participate in this study
- Patients younger than 18 years (legal grounds)
- Patients who do not understand the meaning of the study due to a lack of knowledge of the German language
- Patients with already diagnosed obstructive airway disease
- Patients who smoked on the day of the examination (distortion of the FeNO results and reactivity during BP testing)
- Nitrate-rich meal (e.g. salad) before the examination (false high FeNO values)
- Patients with respiratory infection < 6 weeks before examination (distortion of the FeNO results and/or BP)

Recruitment and taking informed consent

Patients visiting one of the three participating practices of pneumologists will be contacted by a doctoral candidate (BB) from the Medical Faculty of the Technical University of Munich or a Research Associate at the Institute of General Practice and Health Services Research of the Technical University of Munich regarding possible participation in the study. They will check the inclusion and exclusion criteria and will inform the patient about the study. Finally, the attending pneumologist will provide detailed information about the study. In the case of patients who are screened for participation but who ultimately do not participate in the study (due to

disagreement or other reasons), age, gender, and the reason for non-participation will be documented anonymously in order to be able to conduct a non-responder analysis to assess a potential recruitment bias.

Interventions

Patients included in the study will be examined at first contact (time point t1) and 3 months later (time point t2). The diagnostic work-up is summarized in figure 1 and the patient timeline in figure 2.

Initial examination (t1):

During the first presentation for diagnostic work-up in one of the participating practices of pneumologists, patients are examined with a NO-measuring device (NIOX VERO®) as an index test. Afterwards patients will be examined with WBP as a reference standard; a BP test is performed additionally as part of the diagnostic routine, if required by the pneumologist. In addition, patients complete a questionnaire with structured questions about medical history and symptoms. The questionnaire also contains the "Asthma Control Questionnaire (ACQ)" [17]. The ACQ is used to determine the extent of asthma control (controlled, partially controlled, uncontrolled) and the responsiveness to ICS. ICS responsiveness is given if the ACQ score improves by at least 0.5 in the sense of a "minimal important difference" [14, 18]. Type and daily dose of ICS are recorded.

Index test:

The index test is performed with the electrochemically-based NO-measuring device NIOX VERO®. This device is CE-certified, available in national and international markets, and is already widely used in practices. The FeNO measurements are performed once for each patient according to the recommendations of the ATS and ERS [19]. It is a non-invasive measurement since the patient only needs to take a deep breath through the device and exhale evenly. The FeNO measurements are performed by a doctoral candidate, or a research assistant, or a lung function assistant according to the instructions of the manufacturer. FeNO devices and measurements are provided by the Institute of General Practice and Health Services Research of the Technical University of Munich.

Reference test:

Following the FeNO measurement, an examination with WBP is routinely performed and, if required by the pneumologist, BP is performed as part of the diagnostic routine to rule-in or rule-out the diagnosis of asthma. In Germany, both of these assessments are routine tests and would also take place outside the study. Thus, there is no funding of these measurements.

WBP**:** WBP is considered as the reference standard used to diagnose obstructive airway diseases. An obstructive airway disease is indicated if FEV_1 and/or FEV_1/FVC are below their

lower limits of normal [20]. A reversible airway obstruction is diagnosed if the bronchodilation test is positive (ΔFEV_1 > 12% and 200ml). If there is no bronchial obstruction, BP is performed.

Bronchial provocation test (BP): BP is performed to determine bronchial hyperresponsiveness (BHR) to methacholine according to the 1-concentration-4-step dosimeter protocol [21]. This yields similar results as the ATS multi-concentration protocol [22] but offers advantages in clinical practice. The test is considered positive (indicating BHR) if FEV_1 decreases by at least 20% after inhalation of a maximum cumulative methacholine dose of 960µg, and/or if specific airway resistance (sRaw) increases simultaneously by at least 100% and to at least 2.0 kPa*s, and/or if airway resistance (Raw) increases simultaneously by at least 100% and to at least 0.5 kPa*s/L [22, 23].

Follow-up examination after 3 months (t2):

A single BP test as a reference standard for the diagnosis of asthma only reflects the situation at the time of examination. In some cases, patients with a positive BP test do not suffer from asthma (false positive), since the positive predictive value of BP is only about 70% [12, 24]. According to the German guideline “NVL Asthma” and international guideline GINA [1, 2], a minimum of three months of therapy with ICS is recommended at the time of initial diagnosis (maintenance therapy) before a dose reduction can be started (stepping down). Accordingly, after three months, all patients with a positive BP test or with the diagnosis of asthma, respectively, will be asked to return to the practice. During this follow-up appointment patients will receive a FeNO measurement and afterwards they will be examined with WBP. If asthma has been diagnosed at t1 based on BP, BP will be repeated at t2 (if the result of the bodyplethysmographic examination is inconspicuous). ICS-responsiveness is diagnosed when an airway obstruction is reversible or the tolerance to BP increases by at least one level ("doubling dose"). In addition, a potential improvement in respiratory symptoms is assessed by the Asthma Control Questionnaire (ACQ). Potential changes in FeNO values between t1 and t2 will be evaluated exploratory.

Approximately 2% of patients can be diagnosed with false negatives by BP tests (negative predictive value of BP determined in WBP: 98% [12]). Therefore, patients with an inconspicuous BP test are phoned after three months in order to rule out a false negative test result. Patients will be interviewed regarding symptoms and inhaler medication (structured telephone interview). An interview will take about 5 minutes. If patients report persistent respiratory symptoms although the BP test was negative, they will be offered a follow-up examination at the practice of the respective pneumologist. Depending on the findings, another BP test assessed by WBP will be performed. This will be decided by the pneumologist in each individual case.

Diagnostic decision making

A committee of experts (Antonius Schneider, member of the author board of the NVL Asthma; Rudolf A. Jörres, Senior Scientist for Respiratory Diseases, Occupational Medicine, LMU; Konrad Schultz, Medical Director of the Rehabilitation Clinic for Pneumology Bad Reichenhall) will review each diagnosis in consideration of the patient's medical history, WBP investigation, and BP. The respective pneumologists will be contacted in each inconsistent case to clarify the diagnosis. In addition, the committee of experts assesses whether the patients responded to ICS (delayed type of diagnostic study [16]).

At least one criterion must be given for an asthma diagnosis at t1:

1. increase of FEV_1 from baseline by > 12% and by > 200ml during bronchodilation testing if airway obstruction exists (NVL Asthma [1])
2. positive response of FEV_1 or Raw or sRaw during BP test

At least one of the following criteria at t2 must be fulfilled to establish ICS responsiveness at t2:

1. increase of FEV_1 from baseline (t1) by > 12% and by > 200ml (NVL asthma [1])
2. increase of tolerance during BP tests by at least one level
3. improvement by 0.5 score points in the ACQ

If criterion 1 is fulfilled, a BP test is not performed. In addition, it is not performed if the patient reports a worsening of respiratory symptoms since the initial presentation at t1.

Blinding

The FeNO measurements are performed by a doctoral candidate, a research assistant, or a lung function assistant and are documented on a structured sheet. The pneumologist who assesses the results of WBP and BP tests, is blinded to the results of the FeNO measurement. The results of the examinations and the diagnosis made by the pneumologist are documented on a separate sheet.

The committee of experts (who finally diagnoses or excludes asthma in each individual case and assesses whether the patient responded to ICS) is also blinded to the results of FeNO measurement. The committee only has access to the results of bodyplethysmographic measurements, BP tests, and anamnestic information.

Data management and monitoring

Immediately after signing the patient information, consent and data protection declaration, a pseudonymised study ID is assigned to the patient, under which the further data and study results are documented and stored. From now, all other personal data and findings will only be passed on in encrypted form, i.e. neither the name nor the initials nor the exact date of birth

will appear in the encryption code. The patient identification list remains at the Institute of General Practice and Health Services Research and is only accessible to authorized study personnel. The doctoral candidate enters all data from the patient's files, the values of the FeNO measurement, and the values of the lung function tests obtained by WBP in encrypted form into the statistical program SPSS. 5% of the data will be entered twice to estimate the frequency of typing errors. Moreover, all FeNO values and all asthma diagnoses are entered twice to allow a complete correction of possible typing errors. If no more corrections are required in the database, it is closed and will be used for statistical evaluation. The data collection process and the study procedures will be supervised by a research associate who will also perform periodic visits to the practices.

Statistics

Sample size estimation

According to previous studies in practices of pneumologists, a sample size of n=300 can be expected to include about 105 patients with a new diagnosis of asthma. The prevalence in a previous study in a large lung specialist practice was 39% [5]. To be on the safe side, we assume a slightly lower prevalence of 35% for the current multi-center study. The two primary endpoints will be tested confirmatory on two-sided 5% significance levels. A hierarchical test procedure is used to control the global type-1 error at a 5% significance level. Using exact binomial tests, the expected specificity of 95% is first tested against a reference value of 90% assumed under the null hypothesis. If the test result is positive, another confirmatory test of the expected sensitivity of 35% against a reference value of 20% will follow. These tests each achieve a power of 90% with a sample sizes of 195 patients without asthma diagnosis and 105 asthma patients [25]. The total number of patients is therefore 300.

A validation of the diagnostic algorithm (FeNO, "Allergic Rhinitis" and "Wheezing") [10] is performed by means of Wilcoxon (Mann-Whitney) rank sum tests. With the sample sizes mentioned above, this test reaches a power of 80% at a two-sided and exploratory 5% significance level to detect a diagnostic accuracy of AUC = 0.60 [26].

Statistical analysis

Patients participating in the study are characterized by descriptive statistics (mean values, standard deviations, medians, minimum, maximum; absolute and relative frequencies).

As primary and confirmatory analysis, exact binomial tests of sensitivity and specificity at t1 are performed hierarchically at the predetermined FeNO cut-off value of >50ppb, each against a reference value of 90% or 20%, respectively, and at the two-sided 5% significance level. For these measures as well as for PPV and NPV, corresponding 95% confidence intervals are calculated. Fagan nomograms will be provided for the PPV and NPV to enable the exploration

of post-test probabilities depending on the population specific prevalence. The distribution of diagnoses using FeNO and the reference standard will be shown in a cross-table. The reference standard is the diagnosis of asthma made by body plethysmography and bronchoprovocation if necessary. The statistics mentioned above are calculated analogously:

- in the presence of the symptoms "Allergic rhinitis" and "Wheezing" and using a FeNO cut-off value of >33ppb at t1. In addition, the area under the curve (AUC) of the receiver operating characteristic curve (ROC) is determined with a corresponding 95% confidence interval and tested against a reference value of 0.50 using the Wilcoxon (Mann-Whitney) rank sum test.
- for the prediction of ICS responsiveness (determined at t2) using a FeNO cut-off value of >33ppb

In accordance with the secondary hypotheses, exploratory testing of the PPV values will be performed by exact binomial tests on two-sided 5% significance levels against a reference value of 70%.

Changes in the ACQ during follow-up will be estimated by secondary analyses. For this purpose, a composite endpoint related to ICS responsiveness will be developed (at least one out of these criteria must be fulfilled):

1. increase of FEV_1 from baseline (t1) by > 12% and by > 200ml (NVL asthma [1])
2. increase of tolerance during BP tests by one level
3. improvement of 0.5 score points in the ACQ

Moreover, regarding ICS responsiveness potential changes in FeNO values between the first appointment (t1) and the follow-up appointment (t2) will be evaluated exploratory in secondary analyses. Additionally, subgroup analyses related to different cut-off values of bronchial provocation will be performed. Furthermore, the influence of anthropometric parameters on FeNO values will be analysed in secondary analyses.

3. Patient and public involvement

Patients were not involved in the design of this study.

4. Discussion

The present confirmatory diagnostic study aims to prove the diagnostic benefit of FeNO measurement regarding the diagnosis of asthma. FeNO is an attractive diagnostic tool and provides a non-invasive marker of inflammatory processes in the lung [8]. In contrast, BP as the reference standard for diagnosing asthma is time-consuming, cost-intensive, often only available in specific lung function laboratories and bears a small risk of bronchospasm [22]. Therefore, it is reasonable to discuss FeNO measurement as an alternative procedure to

diagnose asthma. Beyond that, there are strong hints that it has added value to determine ICS-responsiveness. Accordingly, a health technology assessment (HTA) found that the inclusion of FeNO measurement into the diagnostic pathway might increase the diagnostic cost-effectiveness [27].

Several studies have already shown a high diagnostic accuracy of FeNO for discerning asthma in patients suspected of suffering from asthma [5, 6]. In most of these studies, values of specificity were superior to those of sensitivity, suggesting that FeNO measurement is more suitable for ruling in than for ruling out the disease [9]. However, a great weakness of the studies published so far is that the optimal FeNO cut-off values were defined post hoc. This probably led to differences when estimating the diagnostic accuracy of FeNO in different studies as well as to discrepancies regarding the optimal cut-off value for diagnosing or excluding asthma. Indeed, it is known that diagnostic algorithms, including cut-off values perform better in the dataset from which they are derived, compared to a dataset with even similar but different individuals [28]. This phenomenon can be explained, amongst other factors, by overfitting, the absence of important predictors, unsatisfactory model derivation, and differences between patient samples [29, 30]. It is therefore essential to validate pre-defined FeNO cut-off values and diagnostic algorithms based on FeNO measurements in a prospective study, e.g. in individuals outside the derivation dataset, in order to be able to determine the adequate place of FeNO measurement in the diagnosis of asthma and in routine care [1, 9]. The present confirmatory study aims to close this gap.

Due to the confirmatory character of this study, three hypotheses are proposed before the study is conducted. Firstly, we hypothesize that a FeNO cut-off value of >50ppb is suitable to diagnose asthma (sensitivity 35%, specificity 95%). Secondly, we test the validity of the assumption that asthma can be diagnosed with a certainty (PPV) of at least 70% at a FeNO value of >33ppb, if the clinical symptoms "allergic rhinitis" and "wheezing" are present. Moreover, in line with the study by Martin et al. [14], we hypothesize that a FeNO value of >33ppb can predict an ICS responsiveness with a certainty (PPV) of at least 70%. We are aware of the discussion about using FeNO measurement better to identify responsiveness to treatment rather than to label patients with a diagnosis [13]. We aim to investigate the diagnostic usefulness regarding these aspects in a confirmatory manner. Thus, the design of the study should be suitable to verify these hypotheses.

The study will be conducted prospectively by enrolling 300 diagnostic-naïve patients from three different practices of pneumologists to increase the generalisability of the study [31]. All patients will be subjected to the reference standard to establish their true diagnosis. In this context, a major strength of the study is that the diagnosis of asthma will be made rigorously on basis of BP in WBP. It has been shown previously that interpretation of BP responsiveness with WBP, including airway resistance, is superior to the interpretation solely based on FEV_1

responsiveness [12]. After 3 months, patients with an asthma diagnosis will be examined again and the diagnosis of asthma will be verified by the expert team in order to ensure the diagnosis, exclude false positive findings, and determine ICS responsiveness. In parallel, patients without an asthma diagnosis will be phoned after three months and asked if respiratory symptoms still persist. Patients with persistent symptoms will be invited for re-evaluation to exclude false negative findings. This procedure enables us to determine the prognostic value of FeNO regarding the diagnosis of asthma, and to compare the diagnostic-prognostic value of FeNO with BP. The diagnosis of each patient, as well as the evaluation of ICS responsiveness, will be made by an expert team based on anamnestic information as well as on lung function measurements, including BP tests. The expert team as well as the pneumologists of the practices are blinded to the results of FeNO measurement to avoid information bias.

A limitation of the study might be that a longer course of disease could be taken into account, e.g. with a 12-month follow-up evaluation. However, this would not allow us to use the optimal time frame of 3 months for determining ICS responsiveness. Moreover, another limitation of the study might be the fact that the presence of allergic rhinitis is reported by the patient without objective validation. However, this represents the typical state of knowledge in clinical practice as it is uncommon to verify the presence of allergic rhinitis with nasal provocation in pneumological practices. Moreover, it has to be mentioned that we could not include special measures to control for adherence regarding ICS inhalation and consequently this aspect cannot be controlled in this study. Beyond that FeNO devices from various manufacturers should be compared since it cannot be excluded that optimal cut-off values differ between devices. We think that determination of FeNO with NIOX VERO will allow a valid estimation, because it measures FeNO at a mouth flow rate of 50 mL/s over ten seconds and a pressure of 10 cm H_2O as per guideline recommendation [32], and NIOX has been used in many diagnostic studies [9]. The present study might be able to enhance the implementation of FeNO in diagnostic guidelines. However, it will not be able to assess the impact of FeNO on diagnostic decision making in routine care and patient outcomes. This point can be only clarified in a clinical impact analysis study, which will be needed in future [28, 33, 34].

5. Ethics and dissemination

The study was approved by the Ethical Committee of the Technical University of Munich (Reference number 122/20 S). Written, informed consent to participate will be obtained from all participants. The study protocol is registered in the German Clinical Trials Register (DRKS00021125, 24 June 2020). The major results of the study will be published in peer-reviewed academic journals and disseminated through conferences.

6. Trial status

Protocol version 1.0. For recruitment the following time frame is planned: First patient in July 2020, last patient in September 2021, last patient out December 2021.

7. List of abbreviations

AUC	area under the curve
BP	bronchial provocation
FeNO	fractional exhaled nitric oxide
FEV_1	forced expiratory volume in one second
FVC	forced vital capacity
ICS	inhaled corticosteroide
NO	nitric oxide
NPV	negative predictive value
PEF	peak expiratory flow
PPV	positive predictive value
Raw	airway resistance
ROC	receiver operating characteristic
sRaw	specific airway resistance
WBP	whole body plethysmography

8. Declarations

Authors' contributions

C.K. prepared the final study protocol, was involved in the development of the design of the study and agreed to be accountable for all aspects of the work. A.H. developed the details of the statistical analysis plan, reviewed the manuscript and commented on drafts of the final manuscript. R.J. helped with the development of the design of the study and with manuscript preparation. K.S. and B.B. contributed to the development of the study protocol and helped with writing. A.S. developed the design of the study and was involved in the development of the statistical analysis plan as well as in manuscript preparation.

Funding statement

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests statement

The authors declare that they have no competing interests.

Acknowledgements

We are grateful to Drs. Hellmann, Wehgartner-Winkler, Faderl, Dankelmann, Vitiello and Schlimok in Augsburg, to Dr. Weber in Starnberg, and to Dr. Powitz in Munich for the opportunity to perform the study in their practices. We also thank the technical personnel for making the study possible through logistic and organizational support. In addition, we appreciate the willingness of all participants to perform the additional FeNO measurements.

9. References

1. Bundesärztekammer (BÄK), Kassenärztliche Bundesvereinigung (KBV), Arbeitsgemeinschaft der Wissenschaftli-chen Medizinischen Fachgesellschaften (AWMF). Nationale VersorgungsLeitlinie Asthma – Langfassung, 4.Auflage. Konsultationsfassung. 2020 [cited: 2020-06-22].www.asthma.versorgungsleitlinien.de
2. Global Initiative for Asthma. Global Strategy for Asthma Management and Prevention, 2020. Available from: www.ginasthma.org
3. Kuenzli N, Stutz EZ, Perruchoud AP, Braendli O, Tschopp J-M, Bolognini G, et al. Peak flow variability in the SAPALDIA study and its validity in screening for asthma-related conditions. American journal of respiratory and critical care medicine. 1999;160:427-34.
4. Tilemann L, Gindner L, Meyer F, Laux G, Szecsenyi J, Schneider A. Diagnostischer Wert der Peak-Flow-Variabilität bei Verdacht auf Asthma bronchiale in der Hausarztpraxis. DMW-Deutsche Medizinische Wochenschrift. 2009;134(41):2053-8.
5. Schneider A, Schwarzbach J, Faderl B, Welker L, Karsch-Volk M, Jorres RA. FENO measurement and sputum analysis for diagnosing asthma in clinical practice. Respir Med. 2013;107(2):209-16.
6. Schneider A, Faderl B, Schwarzbach J, Welker L, Karsch-Volk M, Jorres RA. Prognostic value of bronchial provocation and FENO measurement for asthma diagnosis--results of a delayed type of diagnostic study. Respir Med. 2014;108(1):34-40.
7. Bjermer L, Alving K, Diamant Z, Magnussen H, Pavord I, Piacentini G, et al. Current evidence and future research needs for FeNO measurement in respiratory diseases. Respiratory medicine. 2014;108(6):830-41.
8. Lane C, Knight D, Burgess S, Franklin P, Horak F, Legg J, et al. Epithelial inducible nitric oxide synthase activity is the major determinant of nitric oxide concentration in exhaled breath. Thorax. 2004;59(9):757-60.

9. Karrasch S, Linde K, Rucker G, Sommer H, Karsch-Volk M, Kleijnen J, et al. Accuracy of FENO for diagnosing asthma: a systematic review. Thorax. 2017;72(2):109-16.

10. Schneider A, Linde K, Reitsma JB, Steinhauser S, Rucker G. A novel statistical model for analyzing data of a systematic review generates optimal cutoff values for fractional exhaled nitric oxide for asthma diagnosis. J Clin Epidemiol. 2017;92:69-78.

11. Schneider A, Wagenpfeil G, Jorres RA, Wagenpfeil S. Influence of the practice setting on diagnostic prediction rules using FENO measurement in combination with clinical signs and symptoms of asthma. BMJ Open. 2015;5(11):e009676.

12. Schneider A, Schwarzbach J, Faderl B, Hautmann H, Jorres RA. Whole-Body Plethysmography in Suspected Asthma: A Prospective Study of Its Added Diagnostic Value in 302 Patients. Dtsch Arztebl Int. 2015;112(24):405-11.

13. Alving K. FeNO and suspected asthma: better to identify responsiveness to treatment than to label with a diagnosis. Lancet Respir Med. 2018;6(1):3-5.

14. Martin MJ, Wilson E, Gerrard-Tarpey W, Meakin G, Hearson G, McKeever TM, et al. The utility of exhaled nitric oxide in patients with suspected asthma. Thorax. 2016;71(6):562-4.

15. Price DB, Buhl R, Chan A, Freeman D, Gardener E, Godley C, et al. Fractional exhaled nitric oxide as a predictor of response to inhaled corticosteroids in patients with non-specific respiratory symptoms and insignificant bronchodilator reversibility: a randomised controlled trial. Lancet Respir Med. 2018;6(1):29-39.

16. Knottnerus JA, Muris JW. Assessment of the accuracy of diagnostic tests: the cross-sectional study. J Clin Epidemiol. 2003;56(11):1118-28.

17. Juniper EF, O'Byrne PM, Guyatt GH, Ferrie PJ, King DR. Development and validation of a questionnaire to measure asthma control. Eur Respir J. 1999;14(4):902-7.

18. Juniper EF, Svensson K, Mork AC, Stahl E. Measurement properties and interpretation of three shortened versions of the asthma control questionnaire. Respir Med. 2005;99(5):553-8.

19. Exhaled N. ATS/ERS recommendations for standardized procedures for the online and offline measurement of exhaled lower respiratory nitric oxide and nasal nitric oxide, 2005. Am J Respir Crit Care Med. 2005;171:912-30.

20. Quanjer PH, Stanojevic S, Cole TJ, Baur X, Hall GL, Culver BH, et al. Multi-ethnic reference values for spirometry for the 3–95-yr age range: the global lung function 2012 equations. Eur Respiratory Soc; 2012.

21. Merget R, Jörres RA, Heinze E, Haufs MG, Taeger D, Brüning T. Development of a 1-concentration-4-step dosimeter protocol for methacholine testing. Respiratory medicine. 2009;103(4):607-13.

22. Crapo R. Guidelines for methacholine and exercise challenge testing-1999. This official statement of the American Thoracic Society was adopted by the ATS Board of Directors, July 1999. Am J Respir Crit Care Med. 2000;161:309-29.
23. Criee CP, Sorichter S, Smith HJ, Kardos P, Merget R, Heise D, et al. Body plethysmography--its principles and clinical use. Respir Med. 2011;105(7):959-71.
24. Perpina M, Pellicer C, de Diego A, Compte L, Macian V. Diagnostic value of the bronchial provocation test with methacholine in asthma. A Bayesian analysis approach. Chest. 1993;104(1):149-54.
25. Chow S-C, Wang H, Shao J. Sample size calculations in clinical research: CRC press; 2007.
26. Noether GE. Sample size determination for some common nonparametric tests. Journal of the American Statistical Association. 1987;82(398):645-7.
27. Harnan SE, Tappenden P, Essat M, Gomersall T, Minton J, Wong R, et al. Measurement of exhaled nitric oxide concentration in asthma: a systematic review and economic evaluation of NIOX MINO, NIOX VERO and NObreath. Health technology assessment (Winchester, England). 2015;19(82):1.
28. Moons KG, Kengne AP, Grobbee DE, Royston P, Vergouwe Y, Altman DG, et al. Risk prediction models: II. External validation, model updating, and impact assessment. Heart. 2012;98(9):691-8.
29. Steyerberg EW. Clinical prediction models: Springer; 2019.
30. Toll D, Janssen K, Vergouwe Y, Moons K. Validation, updating and impact of clinical prediction rules: a review. Journal of clinical epidemiology. 2008;61(11):1085-94.
31. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. Annals of internal medicine. 1999;130(6):515-24.
32. Dweik RA, Boggs PB, Erzurum SC, Irvin CG, Leigh MW, Lundberg JO, et al. An official ATS clinical practice guideline: interpretation of exhaled nitric oxide levels (FENO) for clinical applications. American journal of respiratory and critical care medicine. 2011;184(5):602-15.
33. Moons KG, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. Bmj. 2009;338:b606.
34. Cowley LE, Farewell DM, Maguire S, Kemp AM. Methodological standards for the development and evaluation of clinical prediction rules: a review of the literature. Diagnostic and Prognostic Research. 2019;3(1):16.

Legends to figures

Figure 1. Overview of the diagnostic procedure during the conduct of the study.

Figure 2. Standard Protocol Items: Recommendations for interventional trials (SPIRIT) schedule. ACQ, asthma control questionnaire; FeNO, fractional exhaled nitric oxide. *bronchial provocation test is only performed if required by the pneumologist.

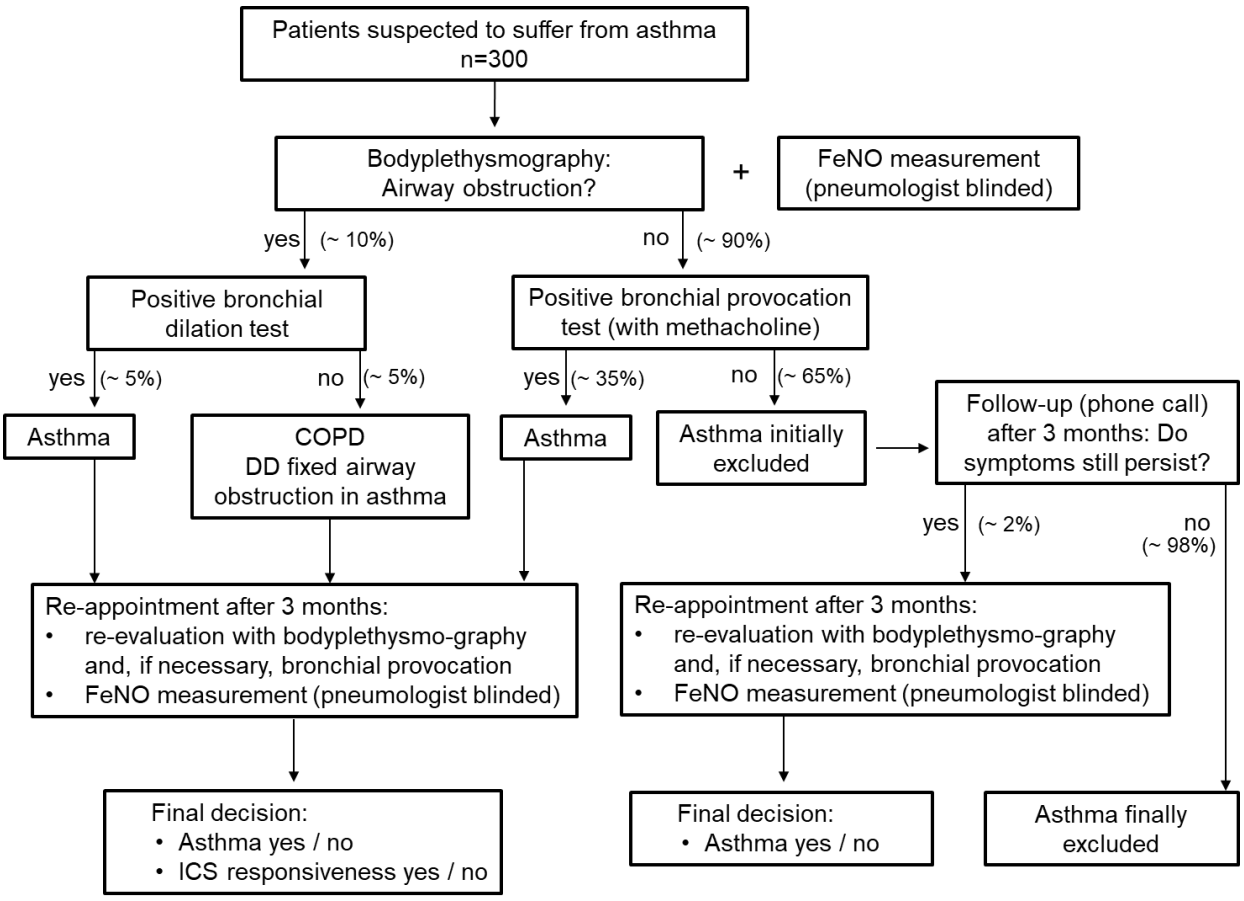


Figure 1. Overview of the diagnostic procedure during the conduct of the study.

	STUDY PERIOD			
	Enrolment	Assessments and interventions	Follow-up (3 months after t1)	
TIMEPOINT	t_0	t_1	t_2	
			Asthma diagnosis at t1	***No asthma diagnosis at t1***
ENROLMENT:				
Eligibility screen	X			
Informed consent	X			
INTERVENTIONS:				
***Index test:* FeNO**		X	X	
Reference test:* Bodyplethysmography and bronchial provocation test		X	X	
ASSESSMENTS:				
ACQ		X	X	
Self-reported questionnaire		X	X	
Structured interview (phone call)				X

Figure 2. Standard Protocol Items: Recommendations for interventional trials (SPIRIT) schedule. ACQ, asthma control questionnaire; FeNO, fractional exhaled nitric oxide. *bronchial provocation test is only performed if required by the pneumologist.

Section & Topic	No	Item	Reported on page #
TITLE OR ABSTRACT			
	1	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)	2
ABSTRACT			
	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	2
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the index test	3-4
	4	Study objectives and hypotheses	4
METHODS			
Study design	**5**	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)	4
Participants	**6**	Eligibility criteria	5
	7	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)	5
	8	Where and when potentially eligible participants were identified (setting, location and dates)	4-5
	9	Whether participants formed a consecutive, random or convenience series	5
Test methods	**10a**	Index test, in sufficient detail to allow replication	6
	10b	Reference standard, in sufficient detail to allow replication	6-7
	11	Rationale for choosing the reference standard (if alternatives exist)	6
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	4
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	6-7
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	8
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	8
Analysis	**14**	Methods for estimating or comparing measures of diagnostic accuracy	9
	15	How indeterminate index test or reference standard results were handled	NA
	16	How missing data on the index test and reference standard were handled	NA
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	9-10
	18	Intended sample size and how it was determined	9
RESULTS			
Participants	**19**	Flow of participants, using a diagram	Figure 1
	20	Baseline demographic and clinical characteristics of participants	NA, we are reporting the study protocol
	21a	Distribution of severity of disease in those with the target condition	NA, we are reporting the study protocol
	21b	Distribution of alternative diagnoses in those without the target condition	NA, we are reporting the study protocol
	22	Time interval and any clinical interventions between index test and reference standard	6-7
Test results	**23**	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	9-10
	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	9-10
	25	Any adverse events from performing the index test or the reference standard	NA
DISCUSSION			
	26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability	10-12
	27	Implications for practice, including the intended use and clinical role of the index test	11-12

OTHER INFORMATION			
	28	Registration number and name of registry	12
	29	Where the full study protocol can be accessed	12
	30	Sources of funding and other support; role of funders	15-16

STARD 2015

AIM

STARD stands for "Standards for Reporting Diagnostic accuracy studies". This list of items was developed to contribute to the completeness and transparency of reporting of diagnostic accuracy studies. Authors can use the list to write informative study reports. Editors and peer-reviewers can use it to evaluate whether the information has been included in manuscripts submitted for publication.

EXPLANATION

A **diagnostic accuracy study** evaluates the ability of one or more medical tests to correctly classify study participants as having a **target condition.** This can be a disease, a disease stage, response or benefit from therapy, or an event or condition in the future. A medical test can be an imaging procedure, a laboratory test, elements from history and physical examination, a combination of these, or any other method for collecting information about the current health status of a patient.

The test whose accuracy is evaluated is called **index test.** A study can evaluate the accuracy of one or more index tests. Evaluating the ability of a medical test to correctly classify patients is typically done by comparing the distribution of the index test results with those of the **reference standard**. The reference standard is the best available method for establishing the presence or absence of the target condition. An accuracy study can rely on one or more reference standards.

If test results are categorized as either positive or negative, the cross tabulation of the index test results against those of the reference standard can be used to estimate the **sensitivity** of the index test (the proportion of participants *with* the target condition who have a positive index test), and its **specificity** (the proportion *without* the target condition who have a negative index test). From this cross tabulation (sometimes referred to as the contingency or "2x2" table), several other accuracy statistics can be estimated, such as the positive and negative **predictive values** of the test. Confidence intervals around estimates of accuracy can then be calculated to quantify the statistical **precision** of the measurements.

If the index test results can take more than two values, categorization of test results as positive or negative requires a **test positivity cut-off**. When multiple such cut-offs can be defined, authors can report a receiver operating characteristic (ROC) curve which graphically represents the combination of sensitivity and specificity for each possible test positivity cut-off. The **area under the ROC curve** informs in a single numerical value about the overall diagnostic accuracy of the index test.

The **intended use** of a medical test can be diagnosis, screening, staging, monitoring, surveillance, prediction or prognosis. The **clinical role** of a test explains its position relative to existing tests in the clinical pathway. A replacement test, for example, replaces an existing test. A triage test is used before an existing test; an add-on test is used after an existing test.

Besides diagnostic accuracy, several other outcomes and statistics may be relevant in the evaluation of medical tests. Medical tests can also be used to classify patients for purposes other than diagnosis, such as staging or prognosis. The STARD list was not explicitly developed for these other outcomes, statistics, and study types, although most STARD items would still apply.

DEVELOPMENT

This STARD list was released in 2015. The 30 items were identified by an international expert group of methodologists, researchers, and editors. The guiding principle in the development of STARD was to select items that, when reported, would help readers to judge the potential for bias in the study, to appraise the applicability of the study findings and the validity of conclusions and recommendations. The list represents an update of the first version, which was published in 2003.

More information can be found on http://www.equator-network.org/reporting-guidelines/stard.