STUDY PROTOCOL
NOVELTY

Prospective observational study in patients with obstructive lung disease: NOVELTY design

Helen K. Reddel[1], Maria Gerhardsson de Verdier[2], Alvar Agustí[3], Gary Anderson[4], Richard Beasley[5], Elisabeth H. Bel[6], Christer Janson[7], Barry Make[8], Richard J. Martin[8], Ian Pavord[9], David Price[10], Christina Keen[11], Asparuh Gardev[12], Stephen Rennard[13], Alecka Sveréus[2], Aruna T. Bansal[14], Lance Brannman[15], Niklas Karlsson[16], Javier Nuevo[17], Fredrik Nyberg[18], Simon S. Young[19] and Jørgen Vestbo[20]

ABSTRACT Asthma and chronic obstructive pulmonary disease (COPD) have overlapping clinical features and share pathobiological mechanisms but are often considered distinct disorders. Prospective, observational studies across asthma, COPD and asthma–COPD overlap are limited. NOVELTY is a global, prospective observational 3-year study enrolling ~12 000 patients ⩾12 years of age from primary and specialist clinical practices in 19 countries (ClinicalTrials.gov identifier: NCT02760329).

NOVELTY's primary objectives are to describe patient characteristics, treatment patterns and disease burden over time, and to identify phenotypes and molecular endotypes associated with differential outcomes over time in patients with a diagnosis/suspected diagnosis of asthma and/or COPD. NOVELTY aims to recruit real-world patients, unlike clinical studies with restrictive inclusion/exclusion criteria.

Data collected at yearly intervals include clinical assessments, spirometry, biospecimens, patient-reported outcomes (PROs) and healthcare utilisation (HCU). PROs and HCU will also be collected 3-monthly *via* internet/telephone. Data will be used to identify phenotypes and endotypes associated with different trajectories for symptom burden, clinical progression or remission and HCU. Results may allow patient classification across obstructive lung disease by clinical outcomes and biomarker profile, rather than by conventional diagnostic labels and severity categories.

NOVELTY will provide a rich data source on obstructive lung disease, to help improve patient outcomes and aid novel drug development.

@ERSpublications
NOVELTY is a global study to characterise patients with asthma and/or COPD and identify novel phenotypes and endotypes http://ow.ly/QFiH30n3IBF

Cite this article as: Reddel HK, Gerhardsson de Verdier M, Agustí A, *et al.* Prospective observational study in patients with obstructive lung disease: NOVELTY design. *ERJ Open Res* 2019; 5: 00036-2018 [https://doi.org/10.1183/23120541.00036-2018].

This article has supplementary material available from openres.ersjournals.com

This study is registered at www.clinicaltrials.gov with identifier number NCT02760329. Data underlying the future findings of the study described in this manuscript will be made available in accordance with AstraZeneca's data sharing policy described at https://astrazenecagrouptrials.pharmacm.com/ST/Submission/Disclosure. The study protocol is available at https://astrazenecagrouptrials.pharmacm.com.

Received: March 03 2018 | Accepted after revision: Nov 13 2018

Affiliations: [1]Woolcock Institute of Medical Research, University of Sydney, Sydney, Australia. [2]Respiratory TA, Global Medical Affairs, AstraZeneca, Mölndal, Sweden. [3]Respiratory Institute, Hospital Clinic, University of Barcelona, IDIBAPS, CIBERES, Barcelona, Spain. [4]Lung Health Research Centre, University of Melbourne, Melbourne, Australia. [5]Medical Research Institute of New Zealand, Wellington, New Zealand. [6]Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands. [7]Dept of Medical Sciences, Uppsala University, Uppsala, Sweden. [8]National Jewish Health and University of Colorado Denver, Denver, CO, USA. [9]Nuffield Dept of Medicine, University of Oxford, Oxford, UK. [10]Observational and Pragmatic Research Institute, Singapore and Centre of Academic Primary Care, University of Aberdeen, Aberdeen, UK. [11]Early Clinical Development IMED Biotech Unit, AstraZeneca, Mölndal, Sweden. [12]Global Medical Affairs, AstraZeneca, Cambridge, UK. [13]Early Clinical Development, IMED Biotech Unit, AstraZeneca, Cambridge, UK. [14]Acclarogen Ltd, Cambridge, UK. [15]US Medical Affairs, AstraZeneca, Gaithersberg, MD, USA. [16]Patient Reported Outcomes, Medical Evidence and Observational Research, Global Medical Affairs, AstraZeneca, Mölndal, Sweden. [17]Respiratory TA, Global Medical Affairs, AstraZeneca, Madrid, Spain. [18]Medical Evidence and Observational Research, Global Medical Affairs, AstraZeneca, Mölndal, Sweden. [19]Precision Medicine and Genomics, IMED Biotech Unit, AstraZeneca, Cambridge, UK. [20]School of Biological Sciences, University of Manchester, Manchester, UK.

Correspondence: Helen K. Reddel, Woolcock Institute of Medical Research, 431 Glebe Point Road, Glebe, NSW 2037, Australia. E-mail: Helen.reddel@sydney.edu.au

Introduction

Asthma and chronic obstructive pulmonary disease (COPD) afflict a combined total of more than 600 million patients globally [1] and are associated with substantial symptom burden, functional impairment and mortality [1, 2]. Although they are often viewed as distinct disorders, both are now acknowledged as heterogeneous and overlapping obstructive lung diseases [3, 4]; an estimated 15–27% of adults with a diagnosis of asthma or COPD have both diagnoses or features of both [5, 6].

Clear evidence exists that patients may have different trajectories in lung function over time [7, 8], with potential genetic and clinical contributors [9], while exacerbation-prone populations have been identified in asthma and COPD [10, 11]. Variations in environmental exposure, disease management and access to care and medicines also contribute to the considerable disparities seen in asthma outcomes worldwide [12]. To improve outcomes, a greater understanding of patient phenotypes (observable characteristics) and endotypes (distinct molecularly defined functional or pathobiological pathways that may be associated with distinct treatment responses) [13] is required across obstructive lung disease and over time (box 1). Biomarkers (defined characteristics measured as an indicator of normal biologic processes, pathogenic processes or responses to an intervention [14]) are key to understanding phenotypes and distinguishing their underlying molecular endotypes, which may help define future therapeutic strategies for distinct patient subgroups.

Although clinical trials are key to defining clinical efficacy and safety of treatments and informing guidelines, regulatory studies exclude up to approximately 95% of patients with asthma or COPD seen in clinical practice [16]. Many studies are limited to patients with only asthma or COPD, with precisely defined disease, and good inhaler technique and adherence, and comorbidities are often an exclusion criterion despite their impact on disease burden and patient outcomes [17–19]. The ecology of care can differ considerably from that in clinical practice [18]. Together, this limits the generalisability of data on which clinical guidelines are largely based.

Recent observational studies provide valuable insights into patient characteristics and associated genetics, biomarkers and phenotypes [4, 20–25]. However, these generally focus on either asthma or COPD [26, 27], are restricted to distinct geographies [28, 29], or study more severe or intensively monitored and treated disease. They may be cross-sectional, so patients following different temporal trajectories [7, 8] may be grouped together. Increasing demand exists for longitudinal observational data in patients across

Box 1. Definitions [13–15]

Phenotype: The observable characteristics of a disease, such as morphology, development, biochemical or physiological properties, or behaviour.

For example, the population with an identified phenotype of obstructive lung disease may share a cluster of clinical, functional and/or inflammatory features, without any implication of a common underlying mechanism.

Endotype: A subtype of disease, defined functionally and pathologically by a distinct molecular mechanism or by distinct treatment responses.

For example, among patients with obstructive lung disease, there are likely to be several specific endotypes associated with divergent underlying molecular causes, and with distinct treatment responses. These endotypes may or may not align with clinical or inflammatory phenotypes identified from studies limited to asthma or to COPD.

Biomarker: A defined characteristic measured as an indicator of normal biologic processes, pathogenic processes or response to an intervention.

obstructive lung disease, including those with comorbidities and from diverse clinical settings, to understand underlying mechanisms.

NOVELTY, a NOVEL observational longiTudinal studY in patients with a diagnosis or suspected diagnosis of asthma and/or COPD, is a large, global, prospective, 3-year cohort study (www.clinicaltrials.gov, NCT02760329; study protocol is available at https://astrazenecagrouptrials.pharmacm.com). Here, we describe the NOVELTY study design and its anticipated outcomes.

Methods

Study design

In NOVELTY, a prospective longitudinal cohort study, patients with a physician diagnosis or suspected diagnosis of asthma and/or COPD will be recruited from clinical practices over 1.5 years, with recruitment stratified by diagnostic label and physician-assessed severity (mild, moderate, severe). Data will be recorded by the treating healthcare professional (HCP) using study-specific electronic case report forms (eCRFs) at yearly visits for 3 years, and by patients every 3 months and in conjunction with yearly visits (figure 1).

Countries participating are Argentina, Australia, Brazil, Canada, mainland China, Colombia, Denmark, France, Germany, Italy, Japan, Mexico, the Netherlands, Norway, South Korea, Spain, Sweden, the UK and the USA.

A range of physicians, including primary care physicians, pulmonologists and allergists, will enrol patients from community and hospital outpatient settings to create a study population drawn from a wide range of asthma and/or COPD severities.

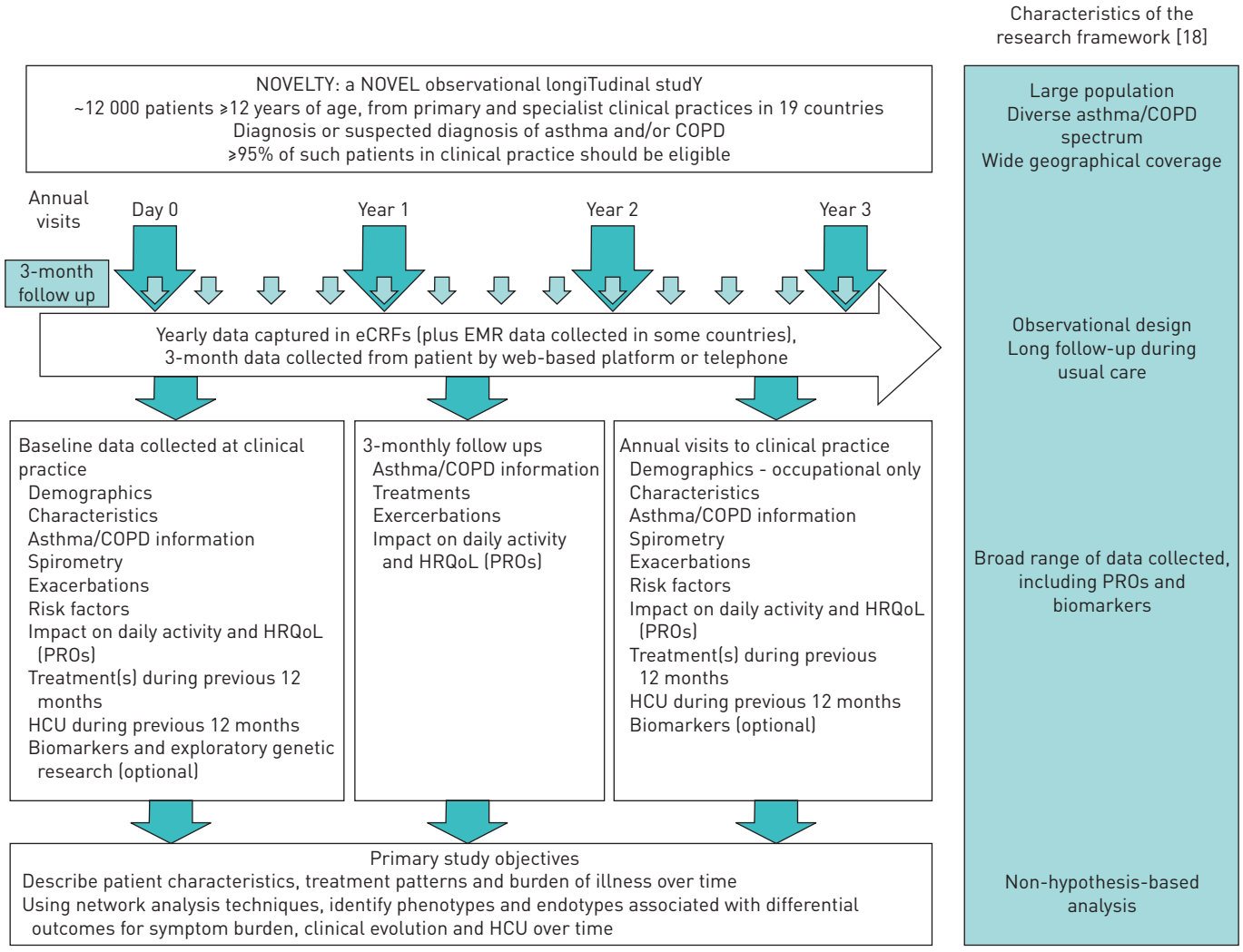

FIGURE 1 NOVELTY study design. COPD: chronic obstructive pulmonary disease; eCRF: electronic case report form; EMR: electronic medical record; HCU: healthcare utilisation; HRQoL: health-related quality of life; PRO: patient-reported outcome.

Study objectives

The primary objectives of NOVELTY:

1) To describe patient characteristics, treatment patterns and burden of illness over time in patients with a diagnosis or suspected diagnosis of asthma and/or COPD.
2) To identify clinical phenotypes and molecular endotypes based on biomarkers and/or clinical parameters that are associated with differential outcomes for symptom burden, clinical evolution and healthcare utilisation (HCU) over time.

The longitudinal analysis will thus focus on differential outcomes in three main domains: symptoms (reflecting burden on patients), clinical evolution (reflecting progression or remission in clinical severity, including lung function) and HCU (reflecting the impact on the health system).

Secondary objectives include comparing patients' diagnostic labels and physician-assessed severity with definitions in existing guidelines and with phenotypic groupings. Patient characteristics, symptom burden, health-related quality of life (HRQoL), exacerbation rates and clinical evolution (*i.e.* clinical progression or remission) will be described in relation to phenotypes and endotypes for subpopulations including patients with recent-onset disease and those with mild or severe disease at enrolment. Specified biomarkers will also be examined to evaluate their stability over time, factors affecting their variability, and their relationships with clinical features and phenotypes.

Other, more exploratory, objectives include descriptions of all-cause and respiratory-related HCU, patient-reported outcomes (PROs; impact on daily activity and HRQoL), and patterns of exacerbations and respiratory infections and their relationship with clinical outcomes. An analysis is planned to assess the adequacy of electronic medical records (EMRs) for obtaining characteristics, clinical progress, and treatment of patients with obstructive lung disease in countries with suitable EMRs (identified from a previous feasibility study conducted in 11 target countries [30]), by comparing them with data collected by NOVELTY eCRFs and PROs.

Patients

Patients will be enrolled from active primary and specialist clinical practices. Patients must have a diagnosis, or a clinically suspected diagnosis, of asthma and/or COPD according to the treating physician; intentionally, no diagnostic criteria are specified. Other inclusion criteria are provision of informed consent (with legal guardian consent for adolescent patients), and age $\geqslant$12 years, although in most countries only patients $\geqslant$18 years of age can be feasibly included. HCPs will be encouraged to include patients with recently diagnosed disease (*i.e.* within 2–5 years before enrolment).

Exclusion criteria are few: participation in an interventional respiratory clinical trial within 12 months prior to NOVELTY enrolment, patients unlikely to complete 3 years of follow up, and patients whose primary respiratory diagnosis (the condition causing most of their respiratory symptoms) is not asthma or COPD. However, co-diagnoses of other respiratory diseases such as bronchiectasis or interstitial lung disease will be accepted. Patients who have received an allogeneic bone marrow transplant or recent whole blood transfusion will be excluded from exploratory genetic research.

Study procedures

Patients will be invited to participate in relation to a routine clinical visit and, once consent is obtained, baseline data will be collected. Patients enrolled during an exacerbation will have baseline data collected 6 weeks later. Patients will have yearly (±3 months) visits with follow-up data collected by their HCP for 3 years or until study discontinuation, whichever comes first. Patients will also complete questionnaires every 3 months (or up to 21 days later than the target date), by a web-based platform or telephone. Patients will receive standard medical care as determined by their physician; no experimental intervention or treatment will be given as part of NOVELTY, and patients will be withdrawn if they enter an interventional trial. Most study procedures (such as history, medications, questionnaires, PROs and optional investigations such as allergy testing and imaging) will be conducted according to the patient's usual standard of care, as per local guidelines. The exceptions are the performance of spirometry, the measurement of fractional exhaled nitric oxide, and the collection, storage and shipment of biosamples, which will be performed in accordance with standardised procedures (see online supplement).

The study will be performed in accordance with ethical principles consistent with the Declaration of Helsinki, International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use and Good Clinical Practice. In each country, the study protocol has been approved by the appropriate Institutional Review Board or Independent Ethics Committee.

Data collection

The variables to be measured are listed in tables 1–3.

TABLE 1 Patient demographics and characteristics to be measured in NOVELTY

Variable	Collection[#]				
	Baseline		Every 3 months	Yearly	
	Physician-reported	Patient-reported	Patient-reported	Physician-reported	Patient-reported
Informed consent	✓				
Demographics					
Age, ethnicity, gender, socioeconomic status, income (USA only), insurance status/ payee (USA only), history of respiratory interventional trials participation	✓				
Occupation	✓			✓	
Characteristics					
Height (post-baseline for adolescents only), weight, BMI, waist circumference, heart rate, pregnancy status	✓			✓	
Risk factors for development of obstructive lung disease					
Smoking status and history, environmental and occupational exposure to pollutants, allergens	✓			✓	

BMI: body mass index; NOVELTY: a NOVEL observational longiTudinal studY. [#]: baseline and yearly data will be collected from healthcare professionals during clinical visits, while 3-monthly data will be collected directly from the patient *via* follow up by web-based platform or telephone, and in conjunction with their yearly clinical visits.

Data will primarily be collected *via* eCRFs completed yearly by the treating HCP, and PROs *via* patient questionnaires every 3 months (with relevant translations). These will provide a consistently collected set of variables aligned to NOVELTY objectives.

The feasibility study revealed that many data required for NOVELTY were not consistently recorded in EMRs [30], necessitating the study-specific eCRF.

Spirometry will be performed by trained site personnel. For sites with suitable spirometers, data will be recorded in the eCRF. For other sites, spirometers (ERT FlowScreen®, ERT, Philadelphia, PA, USA), which meet European Respiratory Society/American Thoracic Society standards [31], will be provided and data collected centrally (centralised over-read performed at baseline only). For predicted values, reference equations from the 2012 Global Lung Function Initiative [32] will be used. Reversibility of airway obstruction will be assessed following withholding of bronchodilators: pre- and post-bronchodilator forced expiratory volume over 1 s and forced vital capacity will be measured immediately before, and ⩾15 min after, a bronchodilator is given. Fractional exhaled nitric oxide will be measured at baseline using Niox Vero devices (Circassia Pharmaceuticals Inc., Morrisville, NC, USA).

PROs will be recorded after the yearly clinic visits and every 3 months (tables 2 and 3). While most respiratory symptom tools are validated for only asthma or COPD, a distinctive feature of NOVELTY is that the same PROs will be administered to the whole study population, irrespective of diagnosis. The chronic airways assessment test (CAAT), a modified version of the COPD assessment test (CAT) [33], will be used with permission of the copyright holder (GlaxoSmithKline, Brentford, UK), with reference to COPD replaced with "your pulmonary disease". As part of CAAT validation, a subset of patients with COPD will also complete the CAT. Generic information on respiratory symptoms will be collected, modelled on guidelines for asthma symptom control, but without referring to asthma. In addition to the above PROs, patients with asthma will complete the Asthma Control Test [34]. The CAPTURE tool (COPD Assessment in Primary Care to Identify Undiagnosed Respiratory Disease and Exacerbation Risk) [35] will also be evaluated in NOVELTY.

Data on medications, comorbidities, exacerbations and HCU will be recorded in the eCRF, and information on medications, exacerbations and HCU will also be collected from patients. Patient-reported adherence will be compared with prescribing data. Information collected on comorbidities includes type, duration and ongoing status.

TABLE 2 Disease-related variables to be measured in NOVELTY

Variable	Collection[#]				
	Baseline		Every 3 months	Yearly	
	Physician-reported	Patient-reported	Patient-reported	Physician-reported	Patient-reported
Asthma/COPD information					
Physician-reported diagnosis and severity	✓				
Age at diagnosis, personal and family history of asthma/COPD and allergies	✓				
Comorbidities	✓			✓	
Asthma/COPD complications, recent respiratory diseases (*e.g.* RTI)	✓	✓	✓	✓	✓
Exacerbations	✓	✓	✓	✓	✓
Symptom assessments: CAAT[¶], RSQ		✓	✓		✓
Symptom assessments: ACT[+], SGRQ		✓			✓
mMRC dyspnoea score	✓				
CAPTURE screening tool		✓			
Lung function measurements (FEV_1, FVC, PEF, $FEF_{25–75\%}$, IC, calculated FEV_1/FVC ratio, calculated FEV_1% predicted)[§]	✓			✓	
Bronchodilator reversibility test	✓				
F_{eNO}	✓				
Biomarkers (with specific consent)[f]					
Blood: differential white blood cell count	✓			✓	
Blood: serum and plasma[##] for biomarker, metabolomic and proteomic analysis	✓			✓	
Blood: DNA and RNA, for genomic and transcriptomic analysis	✓				
Urine sample	✓				

ACT: asthma control test; CAAT: chronic airways assessment test; CAPTURE: COPD Foundation Primary Care Tool for Undiagnosed Respiratory Disease and Exacerbation Risk; COPD: chronic obstructive pulmonary disease; $FEF_{25–75\%}$: forced expiratory flow at 25–75% of the forced vital capacity; F_{eNO}: fractional exhaled nitric oxide; FEV_1: forced expiratory volume over 1 s; FVC: forced vital capacity; IC: inspiratory capacity; mMRC: modified Medical Research Council; NOVELTY: a NOVEL observational longiTudinal studY; PEF: peak expiratory flow; RSQ: Respiratory Symptoms Questionnaire; RTI: respiratory tract infection; SGRQ: St George's Respiratory Questionnaire. [#]: baseline and yearly data will be collected from healthcare professionals during clinical visits, while 3-monthly data will be collected directly from the patient *via* follow up by web-based platform or telephone, and in conjunction with their yearly clinical visits. [¶]: the CAAT is a modified form of the validated COPD assessment test (CAT), which, with consent of the copyright holder (GlaxoSmithKline, Brentford, UK), excludes specific references to COPD in order to also assess patients without a COPD diagnosis; a small subset of patients with COPD will complete both the CAAT and CAT at baseline and year 1, to evaluate the equivalence of the two tests. [+]: only for patients with an asthma diagnosis, for comparison with the CAAT. [§]: for sites where spirometers are provided and data collected centrally, quality control/over-read and Best Test Review will be performed at baseline. [f]: in Brazil, no samples will be collected; in China and Denmark, only samples for haematology will be collected; in Italy, no samples for DNA or RNA analysis will be collected; samples will be collected at each yearly visit in Australia, Canada, France, Germany, Italy, Japan, the Netherlands, Norway, Spain, Sweden, the UK and the USA, and at year 1 only in other countries. [##]: plasma will not be collected from adolescent patients.

Investigators will be asked to record in the eCRF any specialised assessments performed during routine care (*e.g.* diffusion capacity of the lung for carbon monoxide, fractional exhaled nitric oxide, computed tomography scans, 6-minute walk distance and blood clinical chemistry), for which results are available.

With patient consent, blood and urine samples will be collected. Blood samples will be used to measure biomarkers over time, and, with patients' specific consent, for exploratory genetic research, such as investigating genomic, transcriptomic and metabolomic variants associated with the disease phenotypes that may reveal underlying endotypes. Standardised laboratory protocols have been developed for collection,

TABLE 3 Variables relating to the impact of disease, treatment and healthcare utilisation to be collected in NOVELTY

Variable	Collection[#]				
	Baseline		Every 3 months	Yearly	
	Physician-reported	Patient-reported	Patient-reported	Physician-reported	Patient-reported
Impact on daily activity and quality of life					
HRQoL					
SGRQ		✓			✓
EQ-5D-5L		✓	✓		✓
WPAI		✓			✓
Treatment(s) during previous 12 months					
Asthma/COPD treatments	✓	✓	✓	✓	✓
Treatment duration and frequency	✓	✓	✓	✓	✓
Posology	✓			✓	
Patterns of use and treatment adherence for asthma/COPD treatments		✓	✓	✓	✓
Burden of out-of-pocket asthma/COPD treatment expenses[¶]		✓			✓
Reasons for switching/ interruptions/discontinuations of each asthma or COPD treatment		✓	✓	✓	✓
Asthma/COPD treatment satisfaction		✓	✓		✓
Concomitant medications	✓			✓	
Healthcare utilisation during previous 12 months (eCRF) or previous 3 months (PROs)					
Medications, emergency and nonemergency physician visits	✓	✓	✓	✓	✓
Respiratory and nonrespiratory hospitalisations	✓	✓	✓	✓	✓
Out-of-pocket expenses (USA only)		✓			✓
Exacerbations, emergency department visits, hospitalisations, days in ICU	✓	✓	✓	✓	✓
Tests performed during exacerbations	✓			✓	

COPD: chronic obstructive pulmonary disease; eCRF: electronic case report form; EQ-5D-5L: EuroQol 5 dimensions 5 levels health questionnaire; HRQoL: health-related quality of life; ICU: intensive care unit; NOVELTY: a NOVEL observational longiTudinal studY; PRO: patient-reported outcome; SGRQ: St George's Respiratory Questionnaire; WPAI: work productivity and activity impairment. [#]: baseline and yearly data will be collected from healthcare professionals during clinical visits, while 3-monthly data will be collected directly from the patient *via* follow up by web-based platform or telephone, and in conjunction with their yearly clinical visits. [¶]: patients will be asked to consider the burden of out-of-pocket expenses in the 12 months prior to visit.

handling, storage and shipping of biosamples (see online supplement), and all site staff will be trained in sample collection and handling. Biosamples will be stored in a central repository and held for batched analysis (see online supplement). Peripheral blood differential counts will be available for baseline analysis.

In countries selected for EMR analysis, retrospective EMR data may be collected with patient consent, and will be compared with eCRF data to assess EMRs as a source for NOVELTY and future studies, and to evaluate if patients in NOVELTY are representative of nonenrolled patients with a similar diagnosis.

Study organisation and governance

NOVELTY is guided by a scientific committee comprising 12 independent physicians, scientists and statisticians, and seven scientists and economists from the study sponsor. NOVELTY will accept requests for proposals from study investigators, sponsor scientists and the broader respiratory community for

additional analyses and add-on studies (further information will become available on the NOVELTY website, http://aznoveltyproject.com). A governance committee of members from the sponsor and the scientific committee Chairs will be responsible for reviewing proposals.

Sample size calculation

A hypothesis-free approach will be used, focussing on descriptive and exploratory analyses. Therefore, a power calculation for a specified outcome is not strictly relevant; the large overall sample size was chosen to support a wide range of analyses, from general scientific questions to regional or subgroup questions, with sufficient precision. NOVELTY aims to enrol approximately 12 000 patients worldwide; it is assumed that ∼20% of patients with either diagnosis will have both diagnoses [5]. For subgroup analysis of any binary variable or classification with a frequency of 5–95%, the precision with a sample size of 600 would be approximately ±4%; with a sample size of 100, the precision would be approximately ±10%. For data pooled across all countries and severities, the precision would be high even for rarer characteristics (frequency <5% or >95%).

Statistical analysis

To allow for advances in knowledge and statistical methods during the study, analysis will be based on four successive Statistical Analysis Plans. The first will cover baseline data analysis; the second will also include 1-year data and intervening PROs; and the third will include 2-year data and intervening PROs. The fourth Statistical Analysis Plan will include longitudinal study data and analyses to address primary objective 2, which relates to identifying factors contributing to differential outcomes in symptoms, clinical evolution (*i.e.* clinical progression or remission) and HCU over time. Each Statistical Analysis Plan will contain prespecified analyses, and will be finalised prior to database lock for the respective dataset.

Data will be summarised for the overall study population and by prespecified subgroups (including country, demographics, environmental exposures (including tobacco exposure), symptom history, treatment history, concurrent clinical features, treatment setting, socioeconomic setting and healthcare access). For patients who leave the study, data will be censored from the time-point that they leave. Yearly and end-of-study descriptive analyses will be performed for variables in tables 1–3.

Network analysis approaches will be used to identify groups and connections among multiple types and levels of clinical and biological data [36]. Multivariable analysis will be utilised to identify novel phenotypes and endotypes. These analyses will employ exploratory, hypothesis-free techniques including cluster analysis, principal components analysis, data mining, and systems biology analysis. A range of methods will be used to visualise the networks, such as those described previously [37].

Cluster analyses will be performed for both baseline and time-course data using clinical and biomarker features. For example, clinical clusters will be identified on the basis of covariates including symptom burden, features consistent with evolution of clinical severity (worsening or improvement) over time and HCU. The specific measures to be included in the assessment of clinical progression or remission will be pre-specified in the fourth Statistical Analysis Plan to allow for new knowledge emerging by then. Based on current concepts, they may include symptoms, exacerbations, comorbidities (such as obesity and cardiovascular outcomes), and lung function measures.

Internal and external validation

Internal validation will be sought by splitting the baseline data 2:1 into training and validation sets, with the sets balanced for obvious confounders such as age, sex, country of origin and physician diagnosis. This split will be preserved through follow-up years, under the assumption that dropouts should be even across the two groups. Multivariable analyses will be applied to the training set to generate a series of parsimonious prediction models, which may include covariates such as occurrence of exacerbations and other conditions, in relation to clinical outcomes and PROs. Prediction performance will be determined using the validation set. A similar approach will be taken with biomarkers included. The primary analysis of phenotypes and endotypes will be based on longitudinal data, so phenotypes identified in baseline analyses may be superseded by later analyses. The suitability of existing external studies for validation of NOVELTY findings may be limited by the restriction of most of these studies to patients with diagnostic labels of either asthma or COPD, or from one or a few countries, and by use of diagnosis-specific clinical tools and PROs. Because of these differences, external validation of NOVELTY findings may, by necessity, be limited (initially at least) to prespecified physician-diagnosis groups and specific countries of origin.

Discussion

NOVELTY is an ambitious and global research project aiming to go beyond conventional diagnostic labels and existing severity classifications in obstructive lung disease, in order to understand underlying

mechanistic pathways. Asthma and COPD are among the most studied chronic diseases but progress in finding new, more effective treatments has been slow and disappointing compared with other equally complex diseases, where major therapeutic advances and remission-inducing strategies have been made [38]. By recruiting patients across the spectrum of obstructive lung disease, and by using both clinical features and known and emerging biomarkers to identify phenotypes and endotypes associated with differential outcomes, NOVELTY aims to support a paradigm shift towards personalised healthcare.

The strengths of NOVELTY include its size and geographical scope, patient enrolment from active primary and secondary clinical practices (including patients not typically included in clinical or mechanistic trials), inclusion of patients across obstructive lung disease and physician-assigned severities, broad inclusion and minimal exclusion criteria, the wide range of data collected (including samples for biobanking), use of the same parameters across diagnostic labels, long follow up and both retrospective and prospective data collection. By involving 19 countries and approximately 12 000 patients worldwide, NOVELTY will also allow high-precision comparisons between regions beyond the scope of most studies.

It is anticipated that recruiting patients across all disease severities, and patients not normally involved in clinical trials, including those with comorbidities, will result in a study population that samples real-world clinical practice. In contrast to randomised clinical trials that exclude up to approximately 95% of patients [16], ⩾95% of patients ⩾18 years of age with asthma and/or COPD should be eligible for NOVELTY. Stratification of recruitment in each country, aiming for similar numbers across physician-assigned diagnosis/severity categories, should allow sufficient sample sizes for groups such as severe asthma, which constitutes only 3–10% of the general asthma population [39]. By including patients with newly diagnosed disease, there is a greater chance to identify underlying pathobiological mechanisms before the effects of further ageing and environmental influences have been superimposed [40]. Given the global challenges with under-diagnosis of chronic diseases such as asthma and COPD [41], and that NOVELTY aims to describe patient characteristics that could lead to a re-classification of obstructive lung disease, enrolment will not be limited to patients with an established diagnosis but will also include those with respiratory symptoms consistent with either disease, for whom a diagnosis has not yet been made.

In order to identify clinically meaningful phenotypic groups and endotypes associated with differential outcomes, "open" hypothesis-free bioinformatics approaches will be used on longitudinal data. Previously, cluster analyses to define phenotypes have typically used cross-sectional baseline data, then compared outcomes between these clusters [42, 43]. While such analyses will be performed and findings compared with those from large cross-sectional studies such as U-BIOPRED, COPDGene and SPIROMICS, a unique feature of NOVELTY is that we will also identify groups of patients by clinically important outcomes over time (including trajectory of symptoms, clinical evolution (*i.e.* clinical progression or remission) and HCU), then look back to identify any baseline characteristics by which these groups can be predicted.

Limitations of NOVELTY include that patient recruitment *via* clinical practice may bias selection towards those making frequent healthcare visits; HCPs may not enrol patients due to resource or time constraints or be tempted to recruit patients with less complex disease. In addition, physicians may be collecting more information on patient characteristics and lung function than they normally would, which could influence treatment decisions. However, the strategy of comparing patient characteristics from NOVELTY with EMR data from the same practice will allow objective validation and further strengthen the generalisability of the project findings to patients in usual care. As described above, there may initially be limited opportunities for external validation of NOVELTY findings because of a paucity of multinational studies that have recruited patients with asthma and/or COPD, and that have used the same tools regardless of diagnostic label. However, the opportunities for future external validation will be enhanced by publication of the set of NOVELTY tools (including diagnosis-agnostic PROs) so that they can be included in future epidemiologic, clinical and mechanistic studies.

In conclusion, NOVELTY is a very large, worldwide and innovative observational study that goes beyond the scope of many other respiratory observational studies. It is enhanced by broad patient recruitment, a comprehensive longitudinal statistical analysis, PROs, and collection of biospecimens and physiological data for biomarker analysis and molecular endotyping. Results from the study may allow much more precise patient classification according to clinical outcomes and biomarker profiles over time and support the development of novel therapies and a personalised approach across obstructive lung disease. These rich data will augment our understanding of obstructive lung disease, with the ultimate aim of improving patient outcomes.

Acknowledgements: Dirkje Postma, formerly of University of Groningen, University Medical Center, GRAIC Research Institute, Groningen, the Netherlands, provided valuable input into the design of the NOVELTY study, and contributed to earlier drafts of the manuscript. The authors thank Kate Silverthorne and Nina Divorty of CMC CONNECT, a division of Complete Medical Communications Ltd, Manchester, UK, and Glasgow, UK, for providing medical writing support (including development of the first draft under author direction), which was funded by AstraZeneca, Cambridge, UK, in accordance with Good Publication Practice (GPP3) guidelines.

Conflict of interest: H.K. Reddel reports that this study is funded by AstraZeneca. She received reimbursement from AstraZeneca for time spent working on the study design but not for manuscript preparation. Medical writing support was provided by AstraZeneca, as stated in the support statement. She has served on a data and safety monitoring board (DSMB) and advisory boards, and has received unconditional research grants and honoraria from AstraZeneca and GlaxoSmithKline for providing independent medical education, consulting and studying inhalers. She has served on a DSMB for Merck. She has served on a DSMB and an advisory board for, and received an honorarium from Novartis for providing independent medical education. She has received honoraria from Teva and Mundipharma for providing independent medical education. She has received an honorarium from Boehringer Ingelheim for providing independent medical education from and serving on an advisory board. M. Gerhardsson de Verdier is an employee of AstraZeneca. A. Agustí reports receiving personal fees from AstraZeneca during the conduct of the study; and grants and personal fees from AstraZeneca and Menarini, and personal fees from Chiesi, Teva and Novartis, outside the submitted work. G. Anderson has received honoraria from and serves on an advisory board for Novartis; has received honoraria from, serves on an advisory board for and was employed on a sabbatical basis from 2016 to 2017 by AstraZeneca; and has received honoraria and serves on advisory boards for GlaxoSmithKline, Pieris Pharmaceuticals and Boehringer Ingelheim. R. Beasley reports receiving personal fees from Astra Zeneca during the conduct of the study; and grants and personal fees from AstraZeneca, GlaxoSmithKline and the Health Research Council of New Zealand, and grants from Genentech, outside the submitted work. E.H. Bel reports receiving grants and personal fees from Novartis, GSK and AstraZeneca, grants from Roche, personal fees from Sanofi Regeneron, Teva, Vectura and Boehringer, outside the submitted work. C. Janson reports receiving personal fees from AstraZeneca, Boehringer Ingelheim, GlaxoSmithKline, Novartis and Teva, outside the submitted work. Related to the topic of COPD, B. Make reports funding from the NHLBI for the COPDGene study; grants and personal fees from Boehringer Ingelheim, GlaxoSmithKline, AstraZeneca, and Sunovian; personal fees for DSMB from Spiration and Baxalta; CME personal fees from Consensus Medical Education, Integrity Medical Education, WebMD, National Jewish Health, American College of Chest Physicians, Projects in Knowledge, Hybrid Communications, SPIRE Learning, Peer Review Institute, and Medscape; personal fees and other from Mt. Sinai Medical Center; royalties from Up-To-Date; personal fees from Novartis; personal fees from CSL Bering; other from Cleveland Clinic; grants from Pearl (a member of the AstraZeneca Group); and personal fees from Verona, outside the submitted work. R.J. Martin reports receiving personal fees from AstraZeneca for service on a steering committee during the conduct of the study; and travel fees from Respiratory Effectiveness Group, consultancy fees from PMD Healthcare, and grants from MedImmune, Chiesi Farmaceutici SpA and NHLBI, outside the submitted work. I. Pavord reports receiving speaker fees, advisory board honoraria and sponsorship to attend scientific meetings from AstraZeneca, Boehringer Ingelheim and GlaxoSmithKline, a speaker fee from Aerocrine, speaker fees and advisory board honoraria from Almirall and Novartis, advisory board honoraria from Genentech, Regeneron, Merck & Co., Schering-Plough and Mylan Speciality (Dey Pharma) and Respivert, advisory board honoraria and sponsorship to attend scientific meetings from Napp Pharmaceuticals, outside the submitted work. D. Price reports receiving funding from AstraZeneca for the conduct of this study; and board membership (fees paid to Observational and Pragmatic Research Institute) from Aerocrine, Amgen, AstraZeneca, Boehringer Ingelheim, Chiesi, Mylan, Mundipharma, Napp, Novartis, and Teva Pharmaceuticals, consultancy agreements (fees paid to Observational and Pragmatic Research Institute) with Almirall, Amgen, AstraZeneca, Boehringer Ingelheim, Chiesi, GlaxoSmithKline, Mylan, Mundipharma, Napp, Novartis, Pfizer, Teva Pharmaceuticals, and Theravance, grants and unrestricted funding for investigator-initiated studies (conducted through Observational and Pragmatic Research Institute) from Aerocrine, AKL Research and Development Ltd, AstraZeneca, Boehringer Ingelheim, British Lung Foundation, Chiesi, Mylan, Mundipharma, Napp, Novartis, Pfizer, Respiratory Effectiveness Group, Teva Pharmaceuticals, Theravance, UK National Health Service and Zentiva, lectures/speaking engagements (fees paid to Observational and Pragmatic Research Institute) for Almirall, AstraZeneca, Boehringer Ingelheim, Chiesi, CIPLA, GlaxoSmithKline, Kyorin, Mylan, Merck, Mundipharma, Novartis, Pfizer, Skyepharma, and Teva Pharmaceuticals, manuscript preparation (fees paid to Observational and Pragmatic Research Institute) from Mundipharma and Teva Pharmaceuticals, payment for travel/accommodation/meeting expenses (fees paid to Observational and Pragmatic Research Institute) from Aerocrine, AstraZeneca, Boehringer Ingelheim, Mundipharma, Napp, Novartis and Teva Pharmaceuticals, funding for patient enrolment or completion of research (fees paid to Observational and Pragmatic Research Institute) from Chiesi, Novartis, Teva Pharmaceuticals and Zentiva, payment for the development of educational materials (fees paid to Observational and Pragmatic Research Institute) from Mundipharma and Novartis, peer review for grant committees for the Efficacy and Mechanism Evaluation programme and Health Technology Assessment, outside the submitted work; and stock/stock options in AKL Research and Development Ltd, which produces phytopharmaceuticals. He owns 74% of the social enterprise Optimum Patient Care Ltd (Australia, Singapore and UK) and 74% of Observational and Pragmatic Research Institute Pte Ltd (Singapore). C. Keen is an employee of AstraZeneca. A. Gardev is an employee of AstraZeneca. S. Rennard is an employee of AstraZeneca and retains professorship and a part-time appointment at the University of Nebraska Medical Center, Omaha, NE, USA. He reports personal fees from ABIM, Able Associates, Advantage Healthcare, Align2Action, Almirall, APT, ATS, AstraZeneca, Baxter, Boehringer Ingelheim, Cheisi, CIPLA, ClearView Healthcare, the Cleveland Clinic, CME Incite, Complete Medical Group, the COPD Foundation, Cory Paeth, CSA, CSL, CTS Carmel, Dailchi Sankyo, Decision Resources, the Dunn Group, Easton Associates, Elevation Pharma, FirstWord, Forest, the Frankel Group, Gerson, GlaxoSmithKline, Gilead, Grifols, GroupH, Guidepoint Global, Haymarket, HealthStar, Huron Cosulting, Incite, Inthought, IntraMed (Forest), Johnson & Johnson, LEK, McKinsey, Medical Knowledge, MedImmune, the Methodist Health System (Dallas, TX, USA), Navigant, NCI Consulting, Novartis, Nuvis, Pearl (a member of the AstraZeneca Group), Penn Technology, Pfizer, PlanningShop, Prescott, Pro Ed Comm, ProiMed, PSL FirstWord, Pulmatrix, Quadrant, Qwessential, Regeneron, Saatchi and Saatchi, Schlesinger Associates, Strategic North, Synapse, Takeda, Theron, WebMD, NHLBI, the Nebraska DHHS, Otsuka, Pfizer, GlaxoSmithKline, Boehringer Ingelheim, Nycomed, AstraZeneca, Centocor, and Almirall, outside the submitted work. Please note that he has had tobacco industry funding. Specifically, he has received funding from the tobacco industry for studies relating to harm reduction and to the impact of tobacco smoke on stem cells. He has also consulted with R.J. Reynolds, without personal fees, on the topic of harm reduction. He received funding from R.J. Reynolds to evaluate the effect of a harm reduction product in normal smokers (1996) and in subjects with chronic bronchitis (1999) and to assess the effect of smoking cessation on lower respiratory tract inflammation (2000); he participated in a Philip Morris multicentre study to assess biomarkers of smoke exposure (2002); he received funding for a clinical trial from the Institute for Science and Health (2005), which receives support from the tobacco industry, to evaluate biomarkers in exhaled breath associated with

smoking cessation and reduction. This study was supplemented with funding from Lorillard and R.J. Reynolds. He has received a grant from the Philip Morris External Research Program (2005) to assess the impact of cigarette smoking on circulating stem cells in the mouse. He has consulted with R.J. Reynolds on the topic of harm reduction until 2007 but did not receive personal remuneration for this. He has no active tobacco industry-funded projects. All ties with tobacco industry companies and entities supported by tobacco companies were terminated in 2007. A. Sveréus is an employee of AstraZeneca. A.T. Bansal has nothing to disclose. L. Brannman is an employee of AstraZeneca. N. Karlsson is an employee of AstraZeneca. J. Nuevo is an employee of AstraZeneca. F. Nyberg is an employee of AstraZeneca. S.S. Young is an employee of AstraZeneca. J. Vestbo reports receiving an honorarium from AstraZeneca for being a co-principal investigator of the NOVELTY study; and personal fees from GlaxoSmithKline for consultancy for COPD Phase 2 and 3 programme, personal fees from Chiesi Pharmaceuticals, Boeehringer Ingelheim, Novartis and AstraZeneca for consultancy for COPD Phase 2 and 3 programme and payment for lectures including service in speaker bureaus, outside the submitted work.

Support statement: This study is sponsored by AstraZeneca. J. Vestbo is supported by the NIHR Manchester Biomedical Research Centre. Funding information for this article has been deposited with the Crossref Funder Registry.

References

1 Vos T, Flaxman AD, Naghavi M, *et al.* Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012; 380: 2163–2196.
2 Lozano R, Naghavi M, Foreman K, *et al.* Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012; 380: 2095–2128.
3 Moore WC, Fitzpatrick AM, Li X, *et al.* Clinical heterogeneity in the Severe Asthma Research Program. *Ann Am Thorac Soc* 2013; 10: Suppl, S118–S124.
4 Vestbo J, Agusti A, Wouters EFM, *et al.* Should we view chronic obstructive pulmonary disease differently after ECLIPSE? A clinical perspective from the study team. *Am J Respir Crit Care Med* 2014; 189: 1022–1030.
5 Bateman ED, Reddel HK, van Zyl-Smit RN, *et al.* The asthma–COPD overlap syndrome: towards a revised taxonomy of chronic airways diseases? *Lancet Respir Med* 2015; 3: 719–728.
6 Alshabanat A, Zafari Z, Albanyan O, *et al.* Asthma and COPD overlap syndrome (ACOS): a systematic review and meta analysis. *PLoS One* 2015; 10: e0136065.
7 Lange P, Celli B, Agustí A, *et al.* Lung function trajectories leading to chronic obstructive pulmonary disease. *N Engl J Med* 2015; 373: 111–122.
8 McGeachie MJ, Yates KP, Zhou X, *et al.* Patterns of growth and decline in lung function in persistent childhood asthma. *N Engl J Med* 2016; 374: 1842–1852.
9 McGeachie MJ, Yates KP, Zhou X, *et al.* Genetics and genomics of longitudinal lung function patterns in individuals with asthma. *Am J Respir Crit Care Med* 2016; 194: 1465–1474.
10 Hurst JR, Vestbo J, Anzueto A, *et al.* Susceptibility to exacerbation in chronic obstructive pulmonary disease. *N Engl J Med* 2010; 363: 1128–1138.
11 Kupczyk M, ten Brinke A, Sterk PJ, *et al.* Frequent exacerbators – a distinct phenotype of severe asthma. *Clin Exp Allergy* 2014; 44: 212–221.
12 Sastre J, Fabbri LM, Price D, *et al.* Insights, attitudes, and perceptions about asthma and its treatment: a multinational survey of patients from Europe and Canada. *World Allergy Organ J* 2016; 9: 13.
13 Anderson GP. Endotyping asthma: new insights into key pathogenic mechanisms in a complex, heterogeneous disease. *Lancet* 2008; 372: 1107–1119.
14 Biomarkers Definitions Working Group. Biomarkers and surrogate end-points: preferred definitions and conceptual framework. *Clin Pharmacol Ther* 2001; 69: 89–95.
15 Agustí A, Bafadhel M, Beasley R, *et al.* Precision medicine in airway diseases: moving to clinical practice. *Eur Respir J* 2017; 50: 1701655.
16 Herland K, Akselsen JP, Skjonsberg OH, *et al.* How representative are clinical study patients with asthma or COPD for a larger 'real life' population of patients with obstructive lung disease? *Respir Med* 2005; 99: 11–19.
17 Price D, Bateman ED, Chisholm A, *et al.* Complementing the randomized controlled trial evidence base. Evolution not revolution. *Ann Am Thorac Soc* 2014; 11: Suppl. 2, S92–S98.
18 Roche N, Reddel HK, Agustí A, *et al.* Integrating real-life studies in the global therapeutic research framework. *Lancet Respir Med* 2013; 1: e29–e30.
19 Putcha N, Drummond MB, Wise RA, *et al.* Comorbidities and chronic obstructive pulmonary disease: prevalence, influence on outcomes, and management. *Semin Respir Crit Care Med* 2015; 36: 575–591.
20 Faner R, Tal-Singer R, Riley JH, *et al.* Lessons from ECLIPSE: a review of COPD biomarkers. *Thorax* 2014; 69: 666–672.
21 Han MK, Muellerova H, Curren-Everett D, *et al.* GOLD 2011 disease severity classification in the COPDGene cohort. *Lancet Respir Med* 2013; 1: 43–50.
22 Keene JD, Jacobson S, Kechris K, *et al.* Biomarkers predictive of exacerbations in the SPIROMICS and COPDGene cohorts. *Am J Respir Crit Care Med* 2017; 195: 473–481.
23 Kuo CS, Pavlidis S, Loza M, *et al.* A transcriptome-driven analysis of epithelial brushings and bronchial biopsies to define asthma phenotypes in U-BIOPRED. *Am J Respir Crit Care Med* 2017; 195: 443–455.
24 Martinez CH, Diaz AA, Meldrum C, *et al.* Age and small airway imaging abnormalities in subjects with and without airflow obstruction in SPIROMICS. *Am J Respir Crit Care Med* 2017; 195: 464–472.
25 Shaw DE, Sousa AR, Fowler SJ, *et al.* Clinical and inflammatory characteristics of the European U-BIOPRED adult severe asthma cohort. *Eur Respir J* 2015; 46: 1308–1321.
26 SPIROMICS. Participating Institutions. www2.cscc.unc.edu/spiromics/ParticipatingInstitutions Date last updated: 2016. Date last accessed: July 3, 2017.
27 U-BIOPRED. Who is involved? www.europeanlung.org/en/projects-and-research/projects/u-biopred/who-is-involved/ Date last updated: 2016. Date last accessed: July 3, 2017.

28 Aguib Y, Al Suwaidi J. The Copenhagen City Heart Study (Østerbroundersøgelsen). *Glob Cardiol Sci Pract* 2015; 2015: 33.
29 Lifelines. Researchers. https://lifelines.nl/lifelines-research/general Date last updated: 2016. Date last accessed: August 31, 2017.
30 Reddel HK, Gerhardsson de Verdier M, Agustí A, *et al.* An innovative approach to study design: using electronic medical records to inform the feasibility and design of the NOVELTY study (a NOVEL observational longiTudinal studY on patients with asthma and/or COPD). *Thorax* 2016; 71: A157–A158.
31 Miller MR, Hankinson J, Brusasco V, *et al.* Standardisation of spirometry. *Eur Respir J* 2005; 26: 319–338.
32 Quanjer PH, Stanojevic S, Cole TJ, *et al.* Multi-ethnic reference values for spirometry for the 3–95-yr age range: the global lung function 2012 equations. *Eur Respir J* 2012; 40: 1324–1343.
33 Jones PW, Harding G, Berry P, *et al.* Development and first validation of the COPD assessment test. *Eur Respir J* 2009; 34: 648–654.
34 Schatz M, Sorkness CA, Li JT, *et al.* Asthma Control Test: reliability, validity, and responsiveness in patients not previously followed by asthma specialists. *J Allergy Clin Immunol* 2006; 117: 549–556.
35 Martinez FJ, Mannino D, Leidy NK, *et al.* A new approach for identifying patients with undiagnosed chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2017; 195: 748–756.
36 Diez D, Agustí A, Wheelock CE. Network analysis in the investigation of chronic respiratory diseases. From basics to application. *Am J Respir Crit Care Med* 2014; 190: 981–988.
37 Wang B, Mezlini AM, Demir F, *et al.* Similarity network fusion for aggregating data types on a genomic scale. *Nat Methods* 2014; 11: 333–337.
38 Pavord ID, Beasley R, Agustí A, *et al.* After asthma: redefining airways diseases. *Lancet* 2017; 391: 350–400.
39 Hekking PP, Wener RR, Amelink M, *et al.* The prevalence of severe refractory asthma. *J Allergy Clin Immunol* 2015; 135: 896–902.
40 Reddel HK. Treatment of overlapping asthma-chronic obstructive pulmonary disease: can guidelines contribute in an evidence-free zone? *J Allergy Clin Immunol* 2015; 136: 546–552.
41 Falagas ME, Vardakas KZ, Vergidis PI. Under-diagnosis of common chronic diseases: prevalence and impact on human health. *Int J Clin Pract* 2007; 61: 1569–1579.
42 Fingleton J, Travers J, Williams M, *et al.* Treatment responsiveness of phenotypes of symptomatic airways obstruction in adults. *J Allergy Clin Immunol* 2015; 136: 601–609.
43 Bourdin A, Molinari N, Vachier I, *et al.* Prognostic value of cluster analysis of severe asthma phenotypes. *J Allergy Clin Immunol* 2014; 134: 1043–1050.