



CrossMark

What is precision medicine?

Inke R. König^{1,2,3}, Oliver Fuchs^{4,5,6}, Gesine Hansen^{7,8}, Erika von Mutius^{4,5} and Matthias V. Kopp^{2,9}

Affiliations: ¹Institut für Medizinische Biometrie und Statistik, Universität zu Lübeck, Lübeck, Germany. ²German Centre for Cardiovascular Research (DZHK), partner site Hamburg/Kiel/Lübeck, Lübeck, Germany. ³Airway Research Center North (ARCN), Member of the German Center for Lung Research (DZL). ⁴Dr von Hauner Children's Hospital, Ludwig Maximilians University, Munich, Germany. ⁵Comprehensive Pneumology Center Munich (CPC-M), Member of the German Center for Lung Research (DZL). ⁶Dept of Paediatric Respiratory Medicine, Inselspital, University Children's Hospital of Bern, University of Bern, Bern, Switzerland. ⁷Dept of Paediatric Pneumology, Allergology and Neonatology, Hannover Medical School, Hannover, Germany. ⁸Biomedical Research in Endstage and Obstructive Lung Disease Hannover (BREATH), Germany; Member of the German Center for Lung Research (DZL). ⁹Dept of Paediatric Allergy and Pulmonology, Clinic of Paediatrics UKSH, University of Lübeck, Lübeck, Germany.

Correspondence: Inke R. König, Institut für Medizinische Biometrie und Statistik, Universität zu Lübeck, Ratzeburger Allee 160, 23562 Lübeck, Germany. E-mail: inke.koenig@imbs.uni-luebeck.de

 @ERSpublications

Precision medicine is defined as a process involving the derivation of novel taxonomies based on deep phenotyping <http://ow.ly/WHna30dNNOZ>

Cite this article as: König IR, Fuchs O, Hansen G, *et al.* What is precision medicine? *Eur Respir J* 2017; 50: 1700391 [<https://doi.org/10.1183/13993003.00391-2017>].

ABSTRACT The term “precision medicine” has become very popular over recent years, fuelled by scientific as well as political perspectives. Despite its popularity, its exact meaning, and how it is different from other popular terms such as “stratified medicine”, “targeted therapy” or “deep phenotyping” remains unclear. Commonly applied definitions focus on the stratification of patients, sometimes referred to as a novel taxonomy, and this is derived using large-scale data including clinical, lifestyle, genetic and further biomarker information, thus going beyond the classical “signs-and-symptoms” approach.

While these aspects are relevant, this description leaves open a number of questions. For example, when does precision medicine begin? In which way does the stratification of patients translate into better healthcare? And can precision medicine be viewed as the end-point of a novel stratification of patients, as implied, or is it rather a greater whole?

To clarify this, the aim of this paper is to provide a more comprehensive definition that focuses on precision medicine as a process. It will be shown that this proposed framework incorporates the derivation of novel taxonomies and their role in healthcare as part of the cycle, but also covers related terms.

Received: Feb 23 2017 | Accepted after revision: July 15 2017

Support statement: This article was supported by German Research Foundation (DFG; grants #KO2250/5-1 and #KO2250/7-1 to IRK) and by the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung) as part of funding for the German Center for Lung Research (DZL). Funding information for this article has been deposited with the Crossref Funder Registry.

Conflict of interest: Disclosures can be found alongside this article at erj.ersjournals.com

Copyright ©ERS 2017

Introduction

The term “precision medicine” has become very popular in recent years, fuelled by scientific as well as political perspectives [1–6]. It has superseded the term “personalised medicine”, which was defined synonymously [1], but then dismissed with the argument that physicians have always treated patients on a personalised level [7]. Indeed, the personal approach that is an inherent part of the doctor–patient relationship is a central aspect of precision medicine, but is not a new invention. However, new biomedical information might add substantial information beyond signs and symptoms that were previously observable, and the term precision medicine implies the novelty of this concept, which is the incorporation of a wide array of individual data, including clinical, lifestyle, genetic and further biomarker information.

A successful example that is frequently cited in this respect is the determination of the human epidermal growth factor receptor (HER)-2 status in breast cancer patients [8]. Initially, HER-2 was discovered to be a prognostic factor with positive patients having a higher probability of a more aggressive course of disease. Subsequently, clinical trials showed the efficacy of the monoclonal antibody trastuzumab, which is directed against an epitope on the external domain of the HER-2 protein; now, trastuzumab is given only to the subgroup of HER-2-positive females, thus proving the gain of including gene expression data. Examples of cases of obstructive airway diseases are shown in table 1, which summarises phenotyping of individuals based on different investigational approaches.

Despite the obvious success of this and other examples [17], the meaning of precision medicine, and how it is related to or different from other popular terms such as “stratified medicine”, “targeted therapy” or deep phenotyping remains unclear. Commonly used definitions of precision medicine include the following aspects. 1) Focus on result, *i.e.* personalised treatment strategies: some define precision medicine as “treatments targeted to the needs of individual patients on the basis of genetic, biomarker, phenotypic or psychosocial characteristics that distinguish a given patient from other patients with similar clinical presentations” [18]. 2) Focus on process and utilised data: others emphasise the data by describing precision medicine as a model that integrates clinical and other data to stratify patients into novel subgroups; it is hoped that these have a common basis of disease susceptibility and manifestation and thus potentially allow for more precise therapeutic solutions [5, 7]. A similar description is given by the President’s Council of Advisors on Science and Technology: “[...] the tailoring of medical treatment to the individual characteristics of each patient. It [...] [means] the ability to classify individuals into

TABLE 1 Phenotyping of individuals with obstructive lung disease based on different investigational approaches

	Reference	Disease	Variable	Diagnostic and/or prognostic model
Medical history and physical examination	[9]	Asthma	Age and onset of disease	Within-subject response to montelukast is superior to fluticasone in childhood asthma in younger children and children with a shorter disease duration
Lifestyle	[10]	Atopic eczema	Cat exposure and genetics	Filaggrin loss-of-function main mutations (501x and 2282del4) and cat ownership at birth interact in their effects on the development of early-life eczema
Basic laboratory tests	[11–14]	Severe asthma	Eosinophil counts	Eosinophil counts in peripheral blood and/or bronchial lavage are predictors for treatment response to anti-IL-5
Imaging	NA	NA	NA	No data available
Functional diagnostics	[9]	Asthma	Lung function and exhaled NO	Children with asthma respond in a different way to ICS and LTRA using FEV ₁ as a clinical end-point. High NO levels and decreased lung function are predictors of a better treatment response to ICS
Immunology/histology	[15]	Asthma	Cytokine levels	Patients with asthma and high pretreatment levels of serum periostin (surrogate marker of Th2 inflammation) had greater improvement in lung function with the monoclonal anti-IL-13 antibody lebrikizumab than did patients with low periostin levels
Omics	[16]	Asthma	ADRB2	Substitution at position 16 (rs1042713) in ADRB2 is associated with enhanced downregulation and uncoupling of β_2 -receptors. The use of a LABA as an “add-on controller” is associated with increased risk of asthma exacerbation in children carrying one or two A alleles at rs1042713

IL: interleukin; NA: not available; NO: nitric oxide; ICS: inhaled corticosteroid; LTRA: leukotriene receptor antagonist; FEV₁: forced expiratory volume in 1 s; Th2: type 2 helper T-cell; ADRB2: β_2 -adrenoceptor gene; LABA: long-acting β -agonist.

subpopulations that differ in their susceptibility to a particular disease or their response to a specific treatment” [1, 2].

Thus, the focus of commonly applied definitions is on the stratification of patients, sometimes referred to as a novel taxonomy, and this is derived using large-scale data that go beyond the classical “signs-and-symptoms” approach. Finding this novel taxonomy has been described as identifying “treatable traits”, *i.e.* disease subgroups that can be treated in a better way because of more precise and validated phenotypic recognition or due to a better understanding of the critical causal pathways [19].

While these aspects are relevant, this description of precision medicine leaves open a number of questions, such as 1) how can or cannot precision medicine be distinguished from other related terms; 2) when does precision medicine actually begin; 3) what are the target achievements and underlying concepts of the idea of precision medicine; 4) how does the stratification of patients translate into better healthcare; 5) can precision medicine really be viewed as the end-point of a novel stratification of patients, as implied, or is it rather a greater whole.

To clarify this, the aim of this article is to provide a more comprehensive definition that focuses on precision medicine as a process. It will be shown that this proposed framework incorporates the derivation of novel taxonomies and their role in healthcare as part of the cycle, but also covers related terms.

Obviously, precision medicine is not an exclusive need or a unique feature of respiratory medicine. However, asthma and chronic obstructive pulmonary disease (COPD) are the most common chronic diseases worldwide, with increasing prevalence, mortality and morbidity. Specifically, the Global Burden of Disease Study presented rankings for years lived with disability, among which asthma ranked 14th and COPD ranked 5th in 2010. The annual costs of healthcare and lost productivity in the European Union due to COPD are estimated to be €48.4 billion per year and those due to asthma to be €33.9 billion per year [20]. These data highlight the need for optimal management of the most prevalent chronic respiratory diseases, and we therefore focus on asthma and COPD to illustrate the meaning of precision medicine.

Definition of precision medicine as a process

In our framework, precision medicine is defined as a process which is depicted in figure 1 and described further in the following sections. The concept incorporates the following ideas: 1) as the process includes a number of feedback loops, there is no steady end-point of precision medicine where, finally, precise medical care is provided to the patients; 2) the cycle implies that there are ongoing efforts to become ever more precise; 3) finer and more accurate stratifications of patients can be interim results of the overall process, which is captured by the term “stratified medicine”.

An important aspect of this framework therefore is that data assessed in the patients are used to try to develop clinically relevant models, and that the results of these analyses then inform the further assessment of patients, thus emphasising the definition of a process and precision medicine as an evolving result (figure 2).

Deep phenotyping

As a starting point, data are gathered within a suitable group of patients. As a general rule, the patients should be representative of the population of interest, but other specific sampling details depend on the later focus of the study. Examples for this might be as follows. If the foremost aim is to identify novel subgroups of patients (see later), an oversampling of underrepresented cases can be useful; if the focus is on the development of diagnostic or prognostic models, cohort studies recruiting representative samples of patients may be sensible; if the aim is to predict treatment response (see later), different study designs will be required, for instance as described in the literature [21].

Whereas the question of how and which patients are recruited, *i.e.* the study design, therefore depends on the specific research question, the characterising feature in this process is the nature of information gathered in these patients. Specifically, the novel taxonomy aimed for in precision medicine does not rely on the classical signs-and-symptoms approach, but adds data from other sources, such as gene expression analyses.

As an example, asthma is known to be a heterogeneous disease with variation in the degree and type of airway inflammation. To investigate the underlying molecular mechanisms, WOODRUFF *et al.* [22] performed a genome-wide gene expression analysis comparing the expression between subjects with asthma, nonasthmatic smokers and healthy controls. They identified three genes as highly induced in airway epithelial cells from subjects with asthma, namely *POSTN* (periostin), *CLCA2* (chloride channel, calcium-activated, family member 1) and *SERPINB2* (serine peptidase inhibitor, clade B, member 2). Furthermore, interleukin (IL)-13 as a key indicator of T-helper type 2 (Th2) inflammation in asthma was

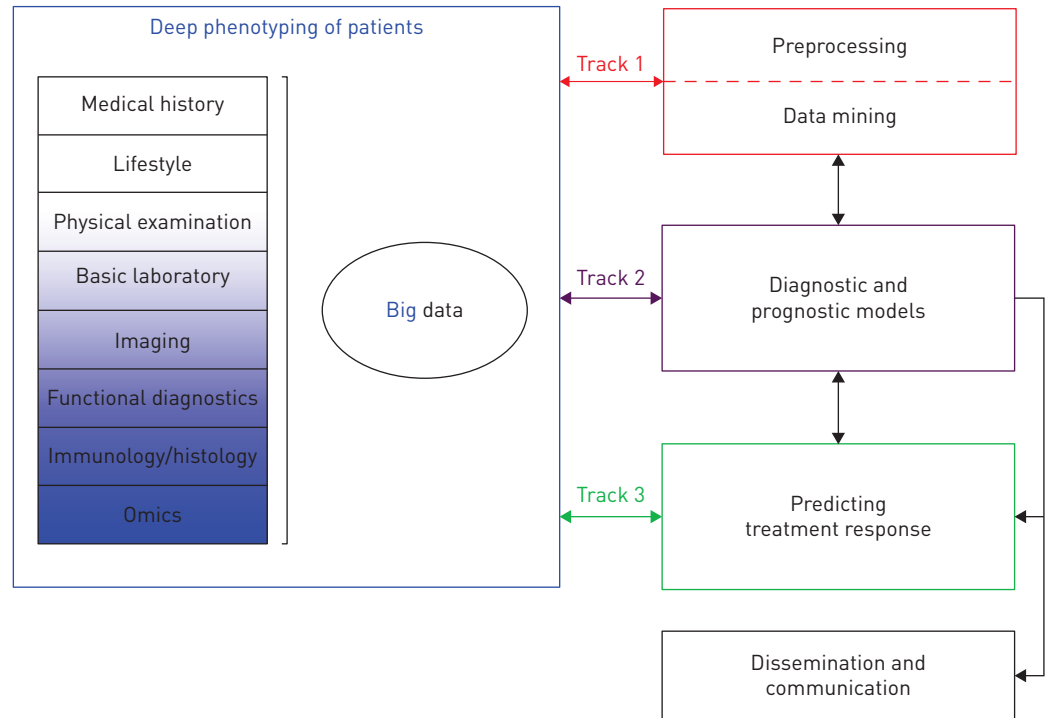


FIGURE 1 The process of precision medicine. In the deep phenotyping stage, information on patients is gathered on different levels. The shading indicates that the more voluminous and complex the dataset becomes, the more likely it is to meet the presupposition for precision medicine and big data. Data are then forwarded for further analysis to tracks 1–3. In track 1, data are preprocessed, including variable selection, and mined for unknown structure. In track 2, variables from the previous stages may be used to develop and validate diagnostic and prognostic models. The clinical relevance of these models may be investigated in studies showing the effect of the implementation of the models or by forwarding the models to track 3. In track 3, specific models are developed and validated which aim at predicting treatment response, partly building on previously developed models. Results from tracks 1–3 are fed back to the deep phenotyping stage to define the subsequent assessment of patients. Models from tracks 2 and 3 need to be disseminated and communicated, providing accessible and easy-to-use algorithms for clinical practice.

found to directly induce the expression of these genes in epithelial cells *in vitro*. The authors postulated that these genes can be viewed as biomarkers of classic IL-13-driven asthma. These data suggest that expression levels in these genes can be considered as part of the data profile of asthma patients.

Further examples illustrating that comprehensive molecular analysis is increasingly being used routinely derive from the Unbiased Biomarkers in Prediction of Respiratory Disease Outcomes (U-BIOPRED) project, which aims to elucidate the mechanisms and biological pathways of severe asthma. In this cohort, Kuo *et al.* [23] identified three transcriptome-associated clusters (TACs) by analysing sputum cell transcriptomics from moderate-to-severe asthmatic subjects and healthy controls. TAC1 comprises an IL-13/Th2-high predominantly eosinophilic cluster, while TAC2 (associated with interferon/tumour necrosis factor- α /inflammasome) and TAC3 are non-Th2 phenotypes. Eosinophilic asthma was predominantly associated with TAC1, but also present in TAC3, whereas neutrophilic asthma was mainly present with TAC2. Of note, previously used biomarkers such as exhaled nitric oxide and periostin were no different in the three TACs. Another example from the U-BIOPRED cohort used transcript profiles for blood gene expression differences and revealed that 1693 genes were differentially expressed between severe asthmatics and nonasthmatics [24]. Cluster analysis of differentially expressed genes defined subgroups among the severe asthmatics that differed in molecular responses to oral corticosteroids. This approach might identify molecular pathways for further studies in poorly controlled asthmatics.

More generally, adding further data with the potential to better reflect underlying pathomechanisms, and therefore relate to prognosis and therapeutic response requires a deep phenotyping, which is the comprehensive and fine-grained description of individuals [5, 25]. Useful illustrations of the different possible levels of assessment have been given by TRACY [26] and LANKTREE *et al.* [27], and the number and depth of layers will increase further over time, for example including data on the set of all expressed phenotypes (*i.e.* phenome-wide) *via* electronic health records [28] or genetic sequence data *via* deep

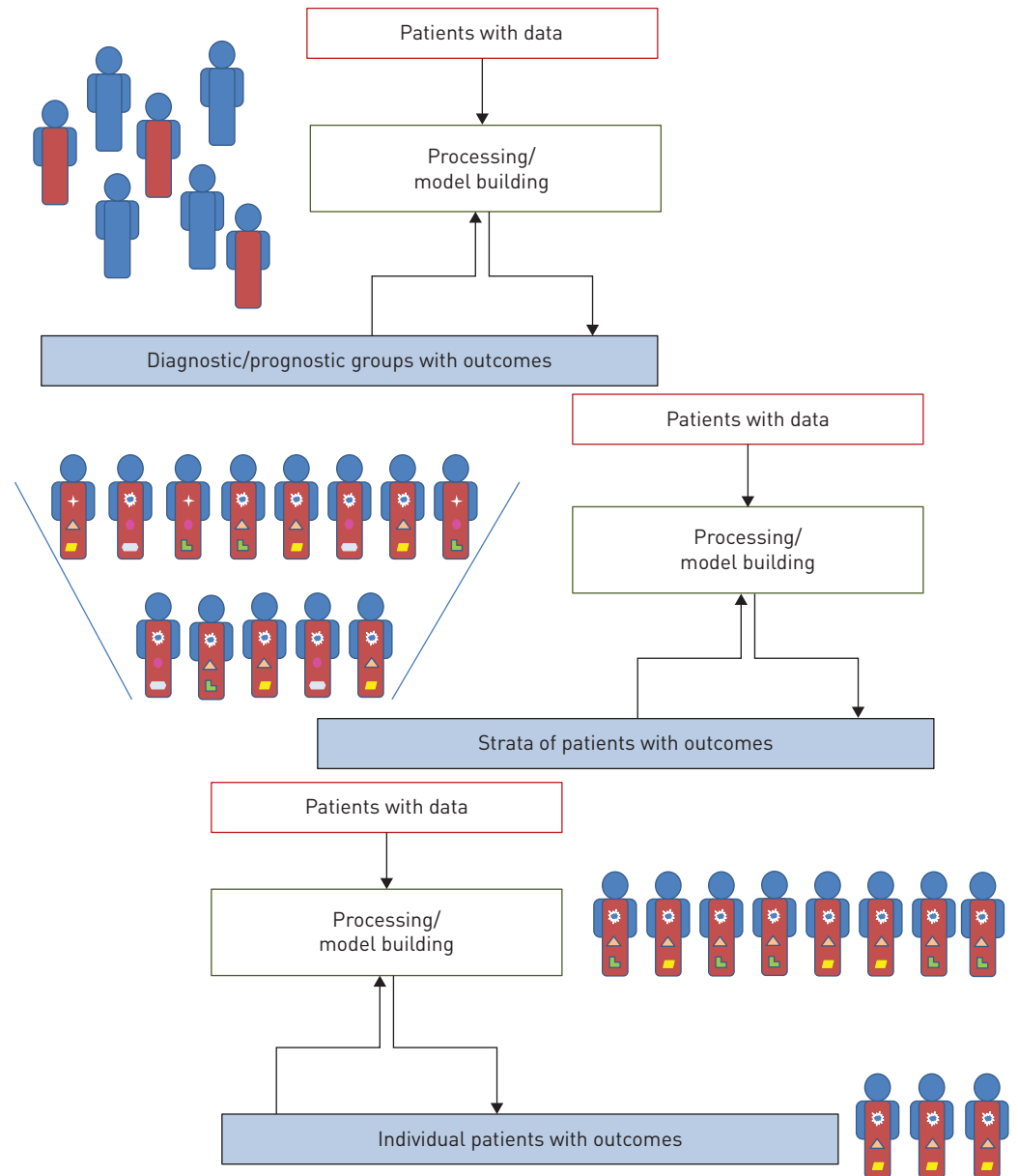


FIGURE 2 Evolving precision medicine. Repeating cycles of patient assessment, data processing and model building yield patient groups at increasingly higher resolution. In the first cycles, patients are categorised into diagnostic and/or prognostic groups based on few obvious characteristics; later cycles define more specific strata of patients using more in-depth data; final cycles may eventually target individual patients with specific data profiles.

sequencing [29] is a realistic option. Longitudinal assessments reflecting disease courses and important points of transition are recognised to be valuable and add further data layers [30].

For instance, the variability and the dynamic of lung function and inflammatory biomarkers over time is an additional challenge, as this might substantially influence the definition of different phenotypes. If physiological and immunological parameters can be assessed on a basis of repeated measurement, this information can be utilised to analyse the temporal pattern in health and disease as well as to disentangle the relationship to the fluctuations of environmental stimuli and clinical symptoms. Numerous specific statistical techniques have been proposed to analyse temporal stability and variability, and overviews can be found in the literature [31–34].

It should be noted that it is difficult to draw an exact line between data sources that indicate what comprises only classical *versus* novel data. Instead, the more voluminous and complex the dataset becomes, the more likely it is to meet the presupposition for precision medicine (indicated by the shading

in figure 1). Depending on the nature and amount of resulting data, it may fulfil the definition of “big data”, regarding the variety of data types, their volume, *i.e.* being too large and complex to be handled with traditional tools, and velocity, *i.e.* having to be processed in a timely manner [35, 36]. However, the data collected here do not necessarily have to be big, but merely surpass the standard examination to foster precision medicine.

Processing deep phenotyping data

How are the large-scale data used for precisifying medicine, *i.e.* integrated and converted into more precise therapeutic interventions [25]? Interestingly, many current reports emphasise the aspect of data storage and handling without acknowledging the role of analytical methods. For instance, a review by SCHADT *et al.* [37] describes the computational environments for the analysis of high-dimensional data, but the only statistical challenge mentioned is the computational complexity associated with these kinds of data. As another example, the Committee on a Framework for Developing a New Taxonomy of Disease [2] postulates a novel informational infrastructure which “requires an ‘information commons’ in which data on large populations of patients become broadly available for research use and a ‘knowledge network’ that adds value to these data by highlighting their inter-connectedness and integrating them with evolving knowledge of fundamental biological processes”. However, how this network “adds value” remains unclear. This is even more important since the novel data types assessed in the previous step require novel approaches from statistics and bioinformatics. For example, the use of next-generation sequencing technologies enables the genome-wide investigation of rare genetic variants, but associating these with diseases requires tailored statistical methods that summarise the information of neighboured variants. And as the generation of sequence data has exploded over the past decade, so has the number of suggested statistical tests [38]. In the more general context of omics data, helpful suggestions on processing have been compiled by McSHANE *et al.* [39] and SUNG *et al.* [40]. Thus, special efforts are required to finally use the generated data in a clinically relevant way.

To further clarify this aspect, we distinguish three partly sequential tracks: in track 1, the data are handled without knowledge of a clinical end-point; in track 2, data are used to build models for a more precise diagnosis or prognosis of disease or disease outcome; and track 3 leads to models that predict more precisely how well specific patients respond to treatment.

It should be stressed that regardless of the chosen track, statistical models are developed that need to be confirmed before further use. Although the literature is heterogeneous with regard to the terminology, we refer to the replication of a model if the model has been shown to be valid in an independent sample drawn from the same population [41, 42] and can be estimated using internal validation techniques such as cross-validation or bootstrapping [43]. In contrast, the validation of a model is shown by testing whether the model is valid in an independent sample drawn from a different population. Depending on the aspects in which the populations differ, this is termed temporal or external validation. Even though internal validation is a crucial first step that can help to optimise a model, temporal or external validation is required for a stringent generalisation of a model [43].

Track 1: preprocessing and data mining

In more detail, track 1 firstly includes quality control and preprocessing of the data. The quality control is strongly dependent on the specific kind of data and therefore not described here; details can be found in the literature [44]. For preprocessing, a number of aspects will be data-type specific, for example, after patients have been genotyped on a genome-wide screening array, allele intensities need to be converted into genotypes by using a genotype calling algorithm [45]; and in a breathomics study, exhaled air is analysed to identify health-related volatile organic compounds. Here, proper preprocessing may include a denoising and baseline correction, and a selection of intensity peaks, among other techniques [46]. Specific quality control considerations are given by McSHANE *et al.* [39].

Other aspects to consider in preprocessing will be more general and may include decisions about how to handle missing data [47, 48]; decisions about how to handle data out of range or below the limits of detection [49, 50]; and harmonisation of corresponding variables in multicentre projects that were assessed or coded slightly differently. For example, studies on fractional exhaled nitric oxide as a marker of eosinophilic inflammation in patients with asthma might be influenced by variable techniques of measurement, sampling procedures, breathing manoeuvres and different types of devices [51].

An important aspect of preprocessing pertains to the preselection of variables. Since the collection of deep phenotyping data usually leads to a great amount of mostly unselected data, this will inevitably include many variables that are irrelevant for later modelling. This is problematic, since many technical examples have shown that if irrelevant variables that just contain random information (so-called “noise”) are added, underlying relationships and structures are masked and not detected [52]. As an example, we consider an

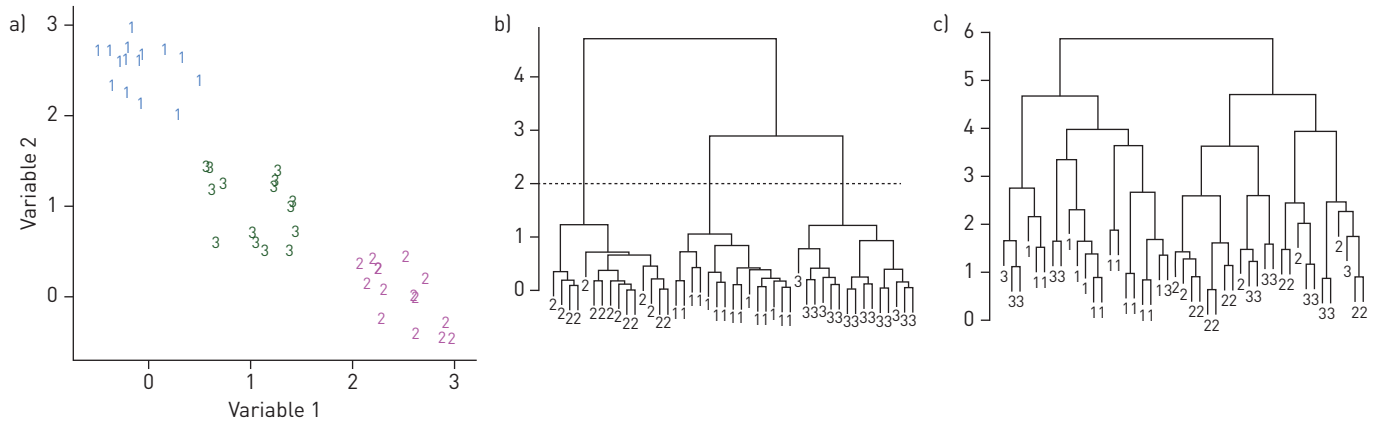


FIGURE 3 Impact of variable selection based on illustration by RAFTERY and DEAN [53]. a) Data of 45 individuals were simulated, who cluster into three homogeneous subgroups using variables 1 and 2. Group membership is indicated by 1, 2 or 3. b) Hierarchical cluster analysis using only these two variables identifies these three clusters perfectly. If clustering is stopped at a distance of ~ 2 , all individuals of groups 1, 2 and 3 fall into distinct clusters. Clustering was based on the Euclidean distance and used complete linkage as agglomeration method. c) Hierarchical cluster analysis using three additional random noise variables fails to identify the three groups. The order in which individuals are grouped is independent of true group membership. Again, clustering was based on the Euclidean distance and used complete linkage as agglomeration method.

artificial situation for which the data of 45 individuals were simulated. These individuals can be clustered into three homogeneous subgroups, as shown in figure 3a if two variables are considered. A hierarchical cluster analysis using only these two variables easily identifies these three clusters perfectly (figure 3b). However, if three random noise variables are added to the dataset, the cluster algorithm fails to find the three groups as indicated in the dendrogram in figure 3c. As a result, almost half of the individuals are classified incorrectly. This “garbage in, garbage out” problem has the rather counterintuitive consequence that having more data is not always better, and if it is planned to derive patient subgroups in the following steps, the removal of noise leads to more stable and better interpretable models. For this, a number of statistical approaches have been suggested that select variables within the clustering procedure. In the example above, a selection within a model-based clustering [53] is able to identify the two relevant variables and discard the others. Furthermore, variable selection might be guided by scientifically based *a priori* knowledge about the possibly biologically relevant variables.

Remaining in track 1, it will in some instances be helpful to interrogate the data for previously unknown structure using data mining techniques. Most prominently, cluster analysis as used in the artificial illustration in figure 3 can be performed to potentially derive meaningful subgroups of individuals, *i.e.* patients with similar profiles based on their similarity in baseline data [54, 55]. If successful and reproducible, these profiles can be viewed as exactly the novel diagnostic taxonomy alluded to in the descriptions of precision medicine given earlier. Since no outcome information would be used, this is an unsupervised learning approach. To continue the previous example on gene expression in asthma, the genetic variables that were found to be promising candidates in the previous step were selected to be taken forward to a subsequent re-analysis of the data including unsupervised learning. Specifically, WOODRUFF *et al.* [56] performed a hierarchical cluster analysis of the gene expression data on the genes *POSTN*, *CLCA2* and *SERPINB2* within individuals with asthma and patients classified per high and low expression, thus interpreted as Th2-high and Th2-low response, respectively.

Subsequently, patients were clustered according to their gene expression levels into those with Th2-high and Th2-low response. Treating asthma patients with the anti-IL-13 antibody lebrikizumab, CORREN *et al.* [15] demonstrated an increase in forced expiratory volume in 1 s (FEV₁) with a distinct effect in Th2-high but not in Th2-low patients, thus demonstrating the clinical utility of this taxonomy.

To fully exploit deep phenotyping data, statistical techniques can be used that consider data from multiple sources simultaneously. For instance, multiple data types such as single nucleotide polymorphisms, copy number variations, methylation states, mRNA expression, protein abundance and clinical information can be clustered in integrative or consensus approaches as suggested in the literature [57, 58].

The result from the preprocessing and data mining in track 1 might be a set of variables and profiles that have possibly already been selected for being good candidates for tracks 2 and 3. Which of these tracks is now chosen depends on the prior knowledge, the specific research aim and the available data. For instance, track 2 is appropriate if representative samples of patients have been assessed longitudinally or

cross-sectionally, bearing the potential to develop or validate prognostic or diagnostic models next. However, if patients are being recruited for a randomised clinical trial, models to predict treatment response may be built directly, thus entering track 3.

Track 2: diagnostic and prognostic models

Using the data from the previous steps, models can be developed that combine multiple predictors to estimate the probability of particular clinical outcomes or events either at the same time (diagnostic model) or in the future (prognostic model). Hence, the main difference between diagnosis and prognosis is the temporal aspect. For instance, 865 adults with severe asthma were recruited prospectively in a cohort study in a closed-panel managed-care organisation to gather risk factors for asthma-related death among adults with severe asthma. Based on these data in a setting with comparable access to care, African-American subjects were not at increased risk of death relative to white subjects. However, in this cohort, greater severity of asthma scores and poorer perceived asthma control scores were indicators of increased mortality risk among adults with severe asthma [59].

A number of aspects have to be considered for model building, such as variable selection, modelling of continuous variables and interactions or restrictions by sample size, and we refer to the literature for details [60–62]. Importantly, out of a multitude of possibilities encompassing classical and machine learning approaches, statistical methods to develop the model have to be chosen [63–65]. This should be context-specific, since some methods automatically include variable selection, some yield easy-to-interpret models, some handle large numbers of variables effortlessly and some automatically take interactions into account. Overall, no method is best, but all have different strengths and weaknesses [39]; furthermore, the classification performances vary depending on the specific dataset [66]. Above that, specific considerations are required if data from multiple sources are to be used to build a model. As described in detail by RITCHIE *et al.* [67], these methods broadly fall into the categories of multi-staged or meta-dimensional analyses. In multi-staged approaches, the analysis is split into a number of steps and performing association analyses using only two data types at a time [68, 69]. In contrast, meta-dimensional techniques combine all data types simultaneously, and examples are provided in the Analysis Tool for Heritable and Environmental Network Associations (ATHENA) methodology [70] or in reference [71].

Once a stable and meaningful model has been derived, the result is an algorithm or a clinical prediction rule [21, 72] that forecasts a clinically relevant outcome such as disease onset or mortality and that is defined by the predicting variables and their respective weights in the algorithm. An example of such an algorithm is a nomogram in which individual values of a patient can be entered to calculate the predicted probability. For example, BERTENS *et al.* [73] present a nomogram that can be used for predicting heart failure, COPD or both combined, based on individual body mass index, pack-years of cigarette smoking, N-terminal pro-brain natriuretic peptide levels, FEV₁, signs of fluid overload, displaced apex beat, abnormal breathing sounds and cough.

These models are always developed to fit the given data in an optimal way. Thus, their actual performance will be estimated too optimistically, and they need to be validated in independent samples [43]. We therefore indicate by the feedback loop in figures 1 and 2 that having developed a promising model instigates further phenotyping of patients in whom the performance of the model can be verified. For example, TANTISIRA *et al.* [74] identified a functional variant of the glucocorticoid-induced transcript 1 gene (GLCCI1) in patients with asthma, which is associated with substantial decrements in the response to inhaled glucocorticoids. Even though they replicated their findings in four independent treatment trials and showed reduced FEV₁ in response to inhaled glucocorticoids in subjects with the variant allele [74], the described variation in the GLCCI1 rs37972 genotype did not seem to affect efficacy of inhaled corticosteroids (ICS) in northern European asthmatic children and young adults. In this population there was no association with an increased risk of either oral corticosteroid use or asthma-related hospital visits or uncontrolled symptoms [75]. The relevance of the step of replication should not be underestimated, meaning that “Indeed, the performance in such a validation study is arguably all that matters, and how a model was derived is of little importance if it performs well” [62].

Furthermore, a prognostic or diagnostic model needs to prove its clinical utility. This means that application of the model helps to improve adverse health outcomes, because effective preventive or interventional measures can be induced based on the results [76, 77]. Without this, the model may still be useful by informing about underlying pathological mechanisms or potential targets for therapeutic approaches, which will be fed back to the previous stages. However, to render medicine more precise, the question of whether implementing the model in practice has the potential to eventually improve patient care needs to be investigated [21]. The previous example is limited in this, because patients with risk-increasing variants still should be treated with ICS. In a meta-analysis by TURNER *et al.* [16], data from 4226 children in five populations were analysed for an interaction between the Arg(16) polymorphism in

the β_2 -adrenoceptor gene (ADRB2) and long-acting β -agonist (LABA) exposure on asthma exacerbations. The authors observed a genetically increased risk for exacerbation only among children treated with ICS plus LABAs, but not after treatment with ICS alone, ICS plus leukotriene receptor antagonist (LTRA) or ICS plus LABAs plus LTRAs. The clinical relevance of this association was shown by a clinical trial in which 62 children with persistent asthma with the homozygous Arg(16) genotype were randomised to receive salmeterol or montelukast as an add-on to inhaled fluticasone [78]. Absence from school was reduced with montelukast compared with salmeterol, and greater improvements occurred in both symptom and quality of life scores with montelukast against salmeterol, thus proving a potential role of Arg(16) genotyping for personalised asthma therapy.

How to prove clinical utility of a given model depends on the context. In some instances, the identified variables may be tested for their value for predicting therapy response in track 3 (see later). However, not all prognostic factors are also predictive of treatment response [79], so that the clinical impact of the model should be assessed directly before claiming its clinical relevance. This means that the effect of using the model on patients' and/or physicians' decision making or directly on participant health outcomes, relative to not using the model, needs to be quantified. Ideally, this is achieved in a cluster-randomised trial as described by MOONS *et al.* [80]. Examples for these studies are scarce, but one illustration is the ongoing progressive cluster-randomised trial in chronic conditions management (PRISMATIC) [81]. In this, practices are randomised to deliver standard care or use the predictive risk stratification tool (Prism) that was developed to estimate risk of an emergency hospital admission in the following year. The clinical utility of this prognostic index is assessed by comparing patients' individual outcomes after 18 months.

It should be emphasised that the development and validation of diagnostic and prognostic models needs to be reported in a transparent way. Recommendations for this have been published with the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) statement [60]. Finally, in order for the models to be used in practice, they need to be communicated and disseminated. For this, scientific publications will usually not suffice, but easy-to-use algorithms have to be made accessible together with interpretation tools.

Track 3: predicting treatment response

To achieve the aim of providing more precise therapeutic strategies, *i.e.*, for “tailored” or “targeted” therapy, variables that define a novel taxonomy need to be assessed with regard to their value in predicting treatment response. Two strategies might guide the process of developing models that predict therapy response. Firstly, one can possibly draw on the prognostic and diagnostic models from the previous stage; in fact, prognostic factors are sometimes assumed to be the natural variables to consider here. For example, epidermal growth factor receptor tyrosine kinase status is a prognostic factor for survival in patients with nonsmall cell lung cancer, but is also predictive for response to the tyrosine kinase inhibitor gefitinib as first-line treatment [82]. Alternatively, biological and other knowledge can be utilised directly from the data gathered, as, for instance, in the evaluation of the monoclonal anti-IL-5 antibody mepolizumab for asthma therapy. In the first clinical trials, mepolizumab was found to be associated with a significant reduction in blood and sputum eosinophils, but did not appear to add significant clinical benefit in patients with asthma [83]. Therefore, subsequent clinical trials selected subjects who had refractory eosinophilic asthma, *i.e.* a sputum eosinophil percentage of $>3\%$ on at least one occasion in the previous 2 years despite high-dose corticosteroid treatment. In this subgroup, mepolizumab therapy showed significant clinical effects, reducing exacerbations and improving asthma quality of life scores [14].

Recently, these findings were expanded for the monoclonal antibody against IL-5 receptor, benralizumab. Again, asthmatics with elevated eosinophils had significantly fewer asthma exacerbations. Moreover, in a *post hoc* analysis of patients who had at least three exacerbations in the previous year, patients with blood eosinophilia receiving benralizumab every 8 weeks had greater improvements relative to placebo in annual exacerbation rate, lung function (FEV₁) and asthma symptoms [84].

As the latter example shows, assessing differential treatment effects often includes the investigation of subgroups in clinical trials to find out whether treatment effects differ meaningfully in different patient groups. While the idea is straightforward, methodological challenges include the definition of clinically relevant differential effects and statistical power [85, 86]. Furthermore, it needs to be acknowledged that if therapeutic decisions should be informed, randomised clinical trials themselves are required, and general aspects that need to be considered in this context have been described in detail in the literature [39]. For these trials, specific study designs have been suggested and reviewed [86, 87]. However, with an increase in precision, these will also become increasingly challenging, because required sample sizes might become difficult to achieve.

Again, track 3 generates further knowledge about disease and treatment, so that there will be another feedback to the phenotyping of patients. An illustration of this is the study on nonsmall cell lung cancer

treated with crizotinib, an inhibitor of the anaplastic lymphoma kinase gene (*ALK*) [88]. In the first part of the trial, a dose-escalation study, patients were identified with lung cancer with *ALK* rearrangement who showed a dramatic improvement after administration of the drug. Based on that finding, the clinical trial was modified, specifically enrolling patients with tumours with this rearrangement. Thus, feeding back the information led to an adjusted trial showing the effectiveness of the treatment.

Finally, as before, in addition to the feedback to phenotyping of patients, dissemination and communication of the taxonomy is a direct and necessary output.

Conclusions

With our process definition, we propose that “precision medicine” should not be viewed as the end-point of a novel stratification of patients with clinical utility, but rather as a highly sophisticated and more complex process. As such, stratified medicine and tailored therapy can be interim outputs of the process, of which the results should be fed back to gain knowledge and thus further increase the precision. Additionally, deep phenotyping of individuals is an integral part of the process, as can be big medical data. In this regard, we would like to stress that just assessing as much data as possible is not expedient; in fact, we need to move beyond the mere hunting and gathering of data and extract those that are relevant. In doing that, prior knowledge should not be confused with mere intuition; the selection should be based on biological, clinical and statistical evidence.

In defining the term precision medicine, it may be helpful to consider what it does not represent. With our process definition, we emphasise that precision medicine is not a steady state or a specific output from the research flow. This implies that “precising medicine” might be the more adequate term to reflect the continuous process of feedback loops, that ongoing efforts for even more precision and individualisation are required and that newly gained knowledge, possibly from new data sources, can yet further improve accuracy (figure 2).

We have discussed the term precision medicine against the background of obstructive airway diseases, *i.e.* bronchial asthma. Obviously, many phenotypes contribute to the so-called asthma syndrome. Considering that the target achievements of precision medicine should be always the most effective treatment for the individual patient, there is an unmet need for ongoing efforts of precising medicine in patients with uncontrolled, severe asthma. However, whether the term precision medicine is applicable and meaningful for all degrees of severity, or even all diseases, is still a matter of debate.

References

- 1 President's Council of Advisors on Science and Technology. Priorities for Personalized Medicine. Washington DC, Executive Office of the President, 2008.
- 2 Committee on a Framework for Developing a New Taxonomy of Disease; National Research Council. Toward Precision Medicine: Building a Knowledge Network for Biomedical Research and a New Taxonomy of Disease. Washington DC, The National Academies Press, 2011.
- 3 Rubin MA. Health: make precision medicine work for cancer care. *Nature* 2015; 520: 290–291.
- 4 Khoury MJ. Planning for the future of epidemiology in the era of big data and precision medicine. *Am J Epidemiol* 2015; 182: 977–979.
- 5 Robinson PN. Deep phenotyping for precision medicine. *Hum Mutat* 2012; 33: 777–780.
- 6 Hamburg MA, Collins FS. The path to personalized medicine. *N Engl J Med* 2010; 363: 301–304.
- 7 McGrath S, Ghera D. Building towards precision medicine: empowering medical professionals for the next revolution. *BMC Med Genomics* 2016; 9: 23.
- 8 Hingorani AD, Windt DA, Riley RD, *et al.* Prognosis research strategy (PROGRESS) 4: stratified medicine research. *BMJ* 2013; 346: e5793.
- 9 Szeffler SJ, Phillips BR, Martinez FD, *et al.* Characterization of within-subject responses to fluticasone and montelukast in childhood asthma. *J Allergy Clin Immunol* 2005; 115: 233–242.
- 10 Bisgaard H, Simpson A, Palmer CN, *et al.* Gene-environment interaction in the onset of eczema in infancy: filaggrin loss-of-function mutations enhanced by neonatal cat exposure. *PLoS Med* 2008; 5: e131.
- 11 Bjermer L, Lemiere C, Maspero J, *et al.* Reslizumab for inadequately controlled asthma with elevated blood eosinophil levels: a randomized phase 3 study. *Chest* 2016; 150: 789–798.
- 12 Corren J, Weinstein S, Janka L, *et al.* Phase 3 study of reslizumab in patients with poorly controlled asthma: effects across a broad range of eosinophil counts. *Chest* 2016; 150: 799–810.
- 13 FitzGerald JM, Bleecker ER, Nair P, *et al.* Benralizumab, an anti-interleukin-5 receptor α monoclonal antibody, as add-on treatment for patients with severe, uncontrolled, eosinophilic asthma (CALIMA): a randomised, double-blind, placebo-controlled phase 3 trial. *Lancet* 2016; 388: 2128–2141.
- 14 Halder P, Brightling CE, Hargadon B, *et al.* Mepolizumab and exacerbations of refractory eosinophilic asthma. *N Engl J Med* 2009; 360: 973–984.
- 15 Corren J, Lemanske RF, Hanania NA, *et al.* Lebrikizumab treatment in adults with asthma. *N Engl J Med* 2011; 365: 1088–1098.
- 16 Turner S, Francis B, Vijverberg S, *et al.* Childhood asthma exacerbations and the Arg16 β 2-receptor polymorphism: a meta-analysis stratified by treatment. *J Allergy Clin Immunol* 2016; 138: 107–113.
- 17 Carneiro BA, Costa R, Taxter T, *et al.* Is personalized medicine here? *Oncology* 2016; 30: 293–303.

- 18 Jameson JL, Longo DL. Precision medicine – personalized, problematic, and promising. *N Engl J Med* 2015; 372: 2229–2234.
- 19 Agusti A, Bel E, Thomas M, *et al.* Treatable traits: toward precision medicine of chronic airway diseases. *Eur Respir J* 2016; 47: 410–419.
- 20 European Respiratory Society. The Burden of Lung Disease. www.erswhitebook.org/chapters/the-burden-of-lung-disease/ Date last accessed: August 25, 2017.
- 21 Wallace E, Smith SM, Perera-Salazar R, *et al.* Framework for the impact analysis and implementation of clinical prediction rules (CPRs). *BMC Med Inform Decis Mak* 2011; 11: 62.
- 22 Woodruff PG, Boushey HA, Dolganov GM, *et al.* Genome-wide profiling identifies epithelial cell genes associated with asthma and with treatment response to corticosteroids. *Proc Natl Acad Sci USA* 2007; 104: 15858–15863.
- 23 Kuo C, Pavlidis S, Loza M, *et al.* T-helper cell type 2 (Th2) and non-Th2 molecular phenotypes of asthma using sputum transcriptomics in U-BIOPRED. *Eur Respir J* 2017; 49: 1602135.
- 24 Bigler J, Boedigheimer M, Schofield J, *et al.* A severe asthma disease signature from gene expression profiling of peripheral blood from U-BIOPRED cohorts. *Am J Respir Crit Care Med* 2017; 195: 1311–1320.
- 25 Delude CM. Deep phenotyping: the details of disease. *Nature* 2015; 527: S14–S15.
- 26 Tracy RP. ‘Deep phenotyping’: characterizing populations in the era of genomics and systems biology. *Curr Opin Lipidol* 2008; 19: 151–157.
- 27 Lanktree MB, Hassell RG, Lahiry P, *et al.* Phenomics: expanding the role of clinical evaluation in genomic studies. *J Investig Med* 2010; 58: 700–706.
- 28 Denny JC, Bastarache L, Roden DM. Phenome-wide association studies as a tool to advance precision medicine. *Annu Rev Genom Hum Genet* 2016; 17: 353–373.
- 29 Le Tourneau C, Kamal M, Tsimberidou AM, *et al.* Treatment algorithms based on tumor molecular profiling: the essence of precision medicine trials. *J Natl Cancer Inst* 2015; 108: djv362.
- 30 Fuchs O, Bahmer T, Rabe KF, *et al.* Asthma transition from childhood into adulthood. *Lancet Respir Med* 2017; 5: 224–234.
- 31 Bravi A, Longtin A, Seely AJ. Review and classification of variability analysis techniques with clinical applications. *Biomed Eng Online* 2011; 10: 90.
- 32 Frey U, Maksym G, Suki B. Temporal complexity in clinical manifestations of lung disease. *J Appl Physiol* 2011; 110: 1723–1731.
- 33 Seely AJ, Macklem PT. Complex systems and the technology of variability analysis. *Crit Care* 2004; 8: R367–R384.
- 34 Thamrin C, Frey U, Kaminsky D, *et al.* Systems biology and clinical practice in respiratory medicine. The twain shall meet. *Am J Respir Crit Care Med* 2016; 194: 1053–1061.
- 35 Chen M, Mao S, Liu Y. Big data: a survey. *Mobile Netw Appl* 2014; 19: 171–209.
- 36 Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. *Health Inf Sci Syst* 2014; 2: 3.
- 37 Schadt EE, Linderman MD, Sorenson J, *et al.* Computational solutions to large-scale data management and analysis. *Nat Rev Genet* 2010; 11: 647–657.
- 38 Dering C, König I, Ramsey L, *et al.* A comprehensive evaluation of collapsing methods using simulated and real data: excellent annotation of functionality and large sample sizes required. *Front Genet* 2014; 5: 323.
- 39 McShane LM, Cavenagh MM, Lively TG, *et al.* Criteria for the use of omics-based predictors in clinical trials: explanation and elaboration. *BMC Med* 2013; 11: 220.
- 40 Sung J, Wang Y, Chandrasekaran S, *et al.* Molecular signatures from omics data: from chaos to consensus. *Biotechnol J* 2012; 7: 946–957.
- 41 Igl B-W, König IR, Ziegler A. What do we mean by ‘replication’ and ‘validation’ in genome-wide association studies? *Hum Hered* 2009; 67: 66–68.
- 42 König IR. Validation in genetic association studies. *Brief Bioinform* 2011; 12: 253–258.
- 43 König IR, Malley JD, Weimar C, *et al.* Practical experiences on the necessity of external validation. *Stat Med* 2007; 26: 5499–5511.
- 44 Ferté C, Trister AD, Huang E, *et al.* Impact of bioinformatic procedures in the development and translation of high-throughput molecular classifiers in oncology. *Clin Cancer Res* 2013; 19: 4315–4325.
- 45 Schillert A, Ziegler A. Genotype calling for the Affymetrix platform. *Methods Mol Biol* 2012; 850: 513–523.
- 46 Smolinska A, Hauschild AC, Fijten RR, *et al.* Current breathomics – a review on data pre-processing techniques and machine learning in metabolomics breath analysis. *J Breath Res* 2014; 8: 027105.
- 47 D’Agostino RB Jr. Overview of missing data techniques. *Methods Mol Biol* 2007; 404: 339–352.
- 48 Graham JW. Missing data analysis: making it work in the real world. *Annu Rev Psychol* 2009; 60: 549–576.
- 49 Lubin JH, Colt JS, Camann D, *et al.* Epidemiologic evaluation of measurement data in the presence of detection limits. *Environ Health Perspect* 2004; 112: 1691–1696.
- 50 Uh HW, Hartgers FC, Yazdanbakhsh M, *et al.* Evaluation of regression methods when immunological measurements are constrained by detection limits. *BMC Immunol* 2008; 9: 59.
- 51 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: 912–930.
- 52 Fowlkes EB, Gnanadesikan R, Kettenring JR. Variable selection in clustering. *J Classif* 1988; 5: 205–228.
- 53 Raftery AE, Dean N. Variable selection for model-based clustering. *J Am Stat Assoc* 2006; 101: 168–178.
- 54 Everitt BS, Landau S, Leese M, *et al.* Cluster Analysis. Chichester, John Wiley & Sons, 2011.
- 55 Milligan GW. Clustering validation: results and implications for applied analyses. In: Arabie P, Hubert LJ, De Soete G, eds. Clustering and Classification. Singapore, World Scientific, 1996.
- 56 Woodruff PG, Modrek B, Choy DF, *et al.* T-helper type 2-driven inflammation defines major subphenotypes of asthma. *Am J Respir Crit Care Med* 2009; 180: 388–395.
- 57 Shen R, Olshen AB, Ladanyi M. Integrative clustering of multiple genomic data types using a joint latent variable model with application to breast and lung cancer subtype analysis. *Bioinformatics* 2009; 25: 2906–2912.
- 58 Kirk P, Griffin JE, Savage RS, *et al.* Bayesian correlated clustering to integrate multiple datasets. *Bioinformatics* 2012; 28: 3290–3297.
- 59 Omachi TA, Iribarren C, Sarkar U, *et al.* Risk factors for death in adults with severe asthma. *Ann Allergy Asthma Immunol* 2008; 101: 130–136.

- 60 Moons KG, Altman DG, Reitsma JB, *et al.* Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* 2015; 162: W1–W73.
- 61 Royston P, Moons KG, Altman DG, *et al.* Prognosis and prognostic research: developing a prognostic model. *BMJ* 2009; 338: b604.
- 62 Steyerberg EW, Moons KG, van der Windt DA, *et al.* Prognosis Research Strategy (PROGRESS) 3: prognostic model research. *PLoS Med* 2013; 10: e1001381.
- 63 König IR, Malley JD, Pajevic S, *et al.* Patient-centered yes/no prognosis using learning machines. *Int J Data Min Bioinform* 2008; 2: 289–341.
- 64 Harrell F. Regression Modeling Strategies. With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis. Cham, Springer International Publishing, 2015.
- 65 Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning. New York, Springer-Verlag, 2009.
- 66 Kruppa J, Liu Y, Diener HC, *et al.* Probability estimation with machine learning methods for dichotomous and multicategory outcome: applications. *Biom J* 2014; 56: 564–583.
- 67 Ritchie MD, Holzinger ER, Li R, *et al.* Methods of integrating data to uncover genotype-phenotype interactions. *Nat Rev Genet* 2015; 16: 85–97.
- 68 Holzinger ER, Ritchie MD. Integrating heterogeneous high-throughput data for meta-dimensional pharmacogenomics and disease-related studies. *Pharmacogenomics* 2012; 13: 213–222.
- 69 Schadt EE, Lamb J, Yang X, *et al.* An integrative genomics approach to infer causal associations between gene expression and disease. *Nat Genet* 2005; 37: 710–717.
- 70 Holzinger ER, Dudek SM, Frase AT, *et al.* ATHENA: the analysis tool for heritable and environmental network associations. *Bioinformatics* 2014; 30: 698–705.
- 71 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.
- 72 Falk G, Fahey T. Clinical prediction rules. *BMJ* 2009; 339: b2899.
- 73 Bertens LC, Moons KG, Rutten FH, *et al.* A nomogram was developed to enhance the use of multinomial logistic regression modeling in diagnostic research. *J Clin Epidemiol* 2016; 71: 51–57.
- 74 Tantisira KG, Lasky-Su J, Harada M, *et al.* Genomewide association between GLCCII and response to glucocorticoid therapy in asthma. *N Engl J Med* 2011; 365: 1173–1183.
- 75 Vijverberg SJ, Tavendale R, Leusink M, *et al.* Pharmacogenetic analysis of GLCCII in three north European pediatric asthma populations with a reported use of inhaled corticosteroids. *Pharmacogenomics* 2014; 15: 799–806.
- 76 Grosse SD, Khoury MJ. What is the clinical utility of genetic testing? *Genet Med* 2006; 8: 448–450.
- 77 Haddow JE, Palomaki GE. A model process for evaluating data on emerging genetic tests. In: Khoury MJ, Little J, Burke W, eds. Human Genome Epidemiology: Scope and Strategies. New York, Oxford University Press, 2004.
- 78 Lipworth BJ, Basu K, Donald HP, *et al.* Tailored second-line therapy in asthmatic children with the Arg(16) genotype. *Clin Sci* 2013; 124: 521–528.
- 79 Clark GM. Prognostic factors *versus* predictive factors: examples from a clinical trial of erlotinib. *Mol Oncol* 2008; 1: 406–412.
- 80 Moons KG, Kengne AP, Grobbee DE, *et al.* Risk prediction models: II. External validation, model updating, and impact assessment. *Heart* 2012; 98: 691–698.
- 81 Hutchings HA, Evans BA, Fitzsimmons D, *et al.* Predictive risk stratification model: a progressive cluster-randomised trial in chronic conditions management (PRISMATIC) research protocol. *Trials* 2013; 14: 301.
- 82 Riley RD, Hayden JA, Steyerberg EW, *et al.* Prognosis Research Strategy (PROGRESS) 2: prognostic factor research. *PLoS Med* 2013; 10: e1001380.
- 83 Flood-Page P, Swenson C, Faiferman I, *et al.* A study to evaluate safety and efficacy of mepolizumab in patients with moderate persistent asthma. *Am J Respir Crit Care Med* 2007; 176: 1062–1071.
- 84 Bleecker ER, FitzGerald JM, Chanez P, *et al.* Efficacy and safety of benralizumab for patients with severe asthma uncontrolled with high-dosage inhaled corticosteroids and long-acting β_2 -agonists (SIROCCO): a randomised, multicentre, placebo-controlled phase 3 trial. *Lancet* 2016; 388: 2115–2127.
- 85 Wang R, Lagakos SW, Ware JH, *et al.* Statistics in medicine – reporting of subgroup analyses in clinical trials. *N Engl J Med* 2007; 357: 2189–2194.
- 86 Ziegler A, Koch A, Krockenberger K, *et al.* Personalized medicine using DNA biomarkers: a review. *Hum Genet* 2012; 131: 1627–1638.
- 87 Buyse M, Michiels S, Sargent DJ, *et al.* Integrating biomarkers in clinical trials. *Expert Rev Mol Diagn* 2011; 11: 171–182.
- 88 Kwak EL, Bang YJ, Camidge DR, *et al.* Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer. *N Engl J Med* 2010; 363: 1693–1703.