Article title: Why the way we define diseases prevents innovation and precision medicine
Authors: Zeinab M. Mamdouh[5,2], Elisa Anastasi[3], Ahmed A. Hassan[5], Johannes Boltze[4], Mahmoud H. Elbatreek[5,2], Cristian Nogales[5], Mayra Pacheco Pachado[5], Alexandra Petraina[5], Zina Piper[5], Alejandro Rodríguez-González[6], Anil Wipat[3], Edwin K. Silverman[7], Ana I Casas[5,8], Emre Guney[9], Harald H.H.W. Schmidt[5]
Affiliations: department of pharmacology & personalised medicine, faculty of health, medicine and life sciences, maastricht university, maastricht, the netherlands[1], department of pharmacology & toxicology, faculty of pharmacy, zagazig university, zagazig, egypt[2], interdisciplinary computing and complex biosystems (icos) group, school of computing, newcastle university, newcastle upon tyne, united kingdom[3], school of life sciences, university of warwick, coventry, united kingdom[4], department of pharmacology & personalised medicine, faculty of health, medicine and life sciences, maastricht university, maastricht, the netherlands[5], centro de tecnología biomédica / ets ingenieros informáticos. universidad politécnica de madrid. pozuelo de alarcón, madrid, spain[6], channing division of network medicine, brigham and women's hospital, boston, massachusetts, usa[7], center for translational neuro- and behavioral sciences (c-tnbs), department of neurology, university hospital essen, essen, germany[8], stalicla, barcelona, catalonia, spain[9]
Orcid ids: 0000-0002-4201-7973[5,2]
Contact e-mail: z.mamdouh@maastrichtuniversity.nl
Journal: DrugRxiv
Review statement: The manuscript is currently under review and should be treated with discretion.
Manuscript submission date: 01 February 2023
Keywords: Network medicine, Systems medicine, Network pharmacology, Endotype, Endophenotype, Taxonomy, Ontology, Classification

Why the way we define diseases prevents innovation and precision medicine

Zeinab M. Mamdouh[1,2], Elisa Anastasi, Ph.D.[3], Ahmed A. Hassan, Ph.D.[1], Johannes Boltze, M.D., Ph.D.[4], Mahmoud H. Elbatreek, Ph.D.[1,2], Cristian Nogales, Ph.D.[1], Mayra Pacheco Pachado[1], Alexandra Petraina[1], Zina Piper[1], Alejandro Rodríguez-González, Ph.D.[5], Anil Wipat, Ph.D.[3], Edwin K. Silverman, M.D. Ph.D.[6], Ana I Casas, Ph.D.[1,7], Emre Guney, Ph.D.[8,*], Harald H.H.W. Schmidt, M.D. Ph.D. Pharm.D.[1,*] For the REPO-TRIAL and REPO4EU consortia

[1] Department of Pharmacology & Personalised Medicine, Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, The Netherlands

[2] Department of Pharmacology & Toxicology, Faculty of Pharmacy, Zagazig University, Zagazig, Egypt

[3] Interdisciplinary Computing and Complex BioSystems (ICOS) Group, School of Computing, Newcastle University, Newcastle upon Tyne, United Kingdom

[4] School of Life Sciences, University of Warwick, Coventry, United Kingdom

[5] Centro de Tecnología Biomédica / ETS Ingenieros Informáticos. Universidad Politécnica de Madrid. Pozuelo de Alarcón, Madrid, Spain

[6] Channing Division of Network Medicine, Brigham and Women's Hospital, Boston, Massachusetts, USA

[7] Center for Translational Neuro- and Behavioral Sciences (C-TNBS), Department of Neurology, University Hospital Essen, Essen, Germany

[8] STALICLA, Barcelona, Catalonia, Spain

Keywords: network medicine; systems medicine; network pharmacology; endotype; endophenotype; taxonomy; ontology; classification

*Authors for correspondence: hschmidt@ppmlab.net, emre.guney@stalicla.com

Abstract

Noncommunicable diseases (NCDs) have become globally abundant, yet the therapeutics we use for them are imprecise. In parallel, identifying new treatments has become more costly than ever due to the ever-aggravating efficacy crisis drug discovery faces. What unites these failures is our ontological classification of diseases, primarily based on descriptive terms. To achieve precision diagnosis and precision therapy in clinical practice, NCDs need to be redefined and subdivided based on their causal molecular mechanisms. However, the inconsistency and incompatibility of the current disease classification systems hinder data integration and analysis towards the characterization of such mechanisms. Here, we explain flaws in the current disease definitions and the dispersion among existing ontologies with the aim of establishing a mechanism-based classification of diseases hence, precision medicine.

Introduction

The current imprecision of therapies leads to a high number of individuals being treated to achieve benefit for at least one patient. Noncommunicable diseases (NCDs) are thus responsible for three-fourths of deaths globally[1]. Moreover, most recently approved compounds that meet regulatory efficacy criteria provide no benefit over existing therapies[2]. The main cause for this imprecision is the knowledge gap between how we define diseases and the mechanisms underlying them. They are often symptom- and organ-based, lacking molecular causal understanding. However, without a corresponding disease mechanism, precision diagnosis and precision therapy are impossible.

Systems medicine aims to overcome symptom- and organ-based silos, by integrating multilevel clinical and omics data and characterising mechanistic relationships between diseases in, so-called, human diseasome (i.e., a network of disease-disease relationships where links are formed by shared risk genes, pleiotropic drugs and clinical comorbidities[3]. Diseases that share risk genes, drugs or pathobiological processes underlying their clinical manifestation across patients cluster together within the diseasome, pointing to common underlying causal biological mechanisms, which can be leveraged for mechanism-based diagnosis and precision therapy. However, the integration of such clinical and omics data is currently prevented by the fact that different biomedical domains, i.e., clinical practice, genetics, and preclinical research, use different and incompatible symptom-, organ-, histology- or other phenotype-based disease classifications[4,5].

Here, we delve into the extent of this problem and suggest that all medical disciplines adopt a single disease ontology, which is then gradually transformed from a descriptive ontology into a mechanistic one. As intriguing as it might sound, the taxonomy of monogenic rare diseases already demonstrates the viability of having such mechanistic ontology. Furthermore, proof-of-concept studies have shown that this approach is also feasible for complex disorders. Revisiting the way we define diseases will enable a conceptually new era of precision medicine and ideally precision disease prevention.

We surrender to our inability to heal

Since 1900, mortality has drastically declined, predominantly through the prevention and treatment of infections[6]. The decline in mortality has not only been stagnating but now it is also experiencing a reversal, and in the USA[7] and UK[8] life expectancy has begun to decrease. Preventive medicine, despite being our sharpest weapon against disease, is grossly underused[9], making drug therapy the most common type of intervention in fighting complex diseases. However, drug therapy is a quite blunt weapon in this fight. Of those drugs that make it to the market, two-thirds fail to provide benefits over existing ones or even induce harm[10], increasing mortality[11,12]. With respect to discovering new drugs, the success rate of the pharmaceutical industry has been on a constant decline[13]. These failures have led the pharma to focus on supposedly easier markets such as immunomodulators and antiviral drugs creating several abandoned disease areas and ultimately a high number of insufficiently treated patients[13]. For a very common indication such as stroke, only a single drug, tPA, is available with more than 30 contraindications, excluding 85% of patients[14].

Moreover, most of the drugs that “work”, do not cure but only alleviate symptoms turning most diseases into chronic conditions. And then, by classifying diseases as chronic, we surrender to our inability to heal. Several shortcomings and false incentives in biomedicine have led to this crisis[5,15–17], but above all, stands our conceptual failure in how we define diseases.

Our definition of diseases is flawed

Unlike rare and infectious diseases, NCDs are not defined by their underlying causal mechanisms (Fig. 1). Instead, NCDs are defined by a symptom linked to a specific organ, such as asthma, hypertension, atherosclerosis, etc., or even, by the name of a clinician, for instance, Alzheimer's or Parkinson's disease. In fact, most diseases are umbrella terms lumping together different causal mechanisms that share this one name-giving phenotype. Therefore, the use of such umbrella disease terms in biomedical research and clinical practice generates an impenetrable mix of molecular mechanisms and clinical comorbidities[18–21]. Our definitions of diseases are often oversimplified generalisations of underlying pathophysiology, making it extremely challenging to develop curative, precise therapies for the entire spectrum of pathological entities under the umbrella term.[22] In some cases, such as syndromes, we summarise medical conditions by a particular group of signs and symptoms but still do not incorporate genetic and molecular perturbations underlying the disease. Consequently, since genetics and clinical medicine seem to live in separate classification systems, taxonomies, or ontologies, the ever-increasing wealth of genetic information does not lead to innovation.

The current heterogeneity in disease classifications is a consequence of our intrinsic desire to bring order to the complexity of clinical medicine. Disease classifications are supposed to fulfil different purposes based on who proposed them, e.g., for diagnoses, subtyping, clinical decision-making, and generating disease models to innovate clinical medicine, etc. As these motives have never been a unified effort, due to historical and practical considerations, a wide array of disease classifications exists, reflecting how diseases are perceived in different domains. Therefore existing classifications differ in various aspects including structure, domain coverage, richness, complexity, community acceptance, maintenance, licence, and construction methods (Table 1 and, for extended information and additional ontology examples, Supplementary Table 1).

Lost in translation: Classification, taxonomy, or ontology?

As we go deeper, it is important to understand the differences between disease classification, taxonomy and ontology, terms that are often confused and used interchangeably. However, they profoundly differ in their richness of information. A 'classification' labels data into categories; a 'taxonomy' provides information where a concept has an *is-a-kind-of* relationship with a broader term; and an 'ontology' enriches data with information on the relationships of defined terms with each other by using a subclass hierarchy. An ontology allows the addition of complex features, the application of different vocabularies on the elements' definitions, and the introduction of logic-based axioms. All provide structural concepts, but an ontology provides further information about the concepts and their relationships[23]. To avoid confusion between these terms, here, we use classification and taxonomy inclusively to refer to different means of disease definition and categorisation systems including ontologies, nomenclatures, dictionaries and vocabularies. In general, four discernible groups of disease classifications can be identified by purpose:

(i) Clinical such as International Classification of Diseases (ICD)[24] and Systematized Nomenclature of Medicine – Clinical Terms (SNOMED CT)[4]

(ii) Biomedical research such as Online Mendelian Inheritance in Man (OMIM for genetic phenotypes)[25], and Medical Subject Headings (MeSH)[26]

(iii) Specific domains such as cardiovascular disease ontology (CVDO)[27] and Orphanet[28] for rare diseases

(iv) Consolidation and integration initiatives such as Disease Ontology (DO)[29], OBO-Foundry[30], DisMaNET[31], Mondo[4,32], and Experimental Factor Ontology (EFO)[33].

The ICD, the widely accepted standard for diagnosis and inpatient hospital coding in the USA and some European countries, was initially developed in 1893 and has been recurrently updated by the World Health Organization[24]. Its codes help to unify electronic health records to facilitate medical claims and to enable data aggregation and clinical studies on specific diseases[34]. ICD supports the operational structure of medicine in organ-based departments (i.e., cardiology, neurology) and vocational specialisations (i.e., cardiologist, neurologist). However, ICD definitions lack specificity induced by variation of code assignment, which hinders data extraction and analysis[35–38]. MeSH[26] , on the other hand, indexes and annotates medical literature through standardised terms referring to defined concepts and hierarchies. Accordingly, it is frequently used in basic and preclinical research as well as biomedical literature mining via popular resources such as MEDLINE or clinicaltrials.gov[39–41]. However, MeSH also presents drawbacks, being inadequate in coverage as the annotation process is subjective, possibly introducing inconsistencies[42,43]. Clinicians and biomedical researchers, all use different disease terms and ontologies, creating a disconnect between clinical research and clinical practice.

Attempts to harmonise disease definitions have failed

To address shortcomings such as structural organisation, classification specificity, or direct applicability, an increasing number of general classification systems have been suggested. DO[29] was developed to be less broad than MeSH and focus only on disease concepts, aiming to harmonise different taxonomies to enable a unified disease annotation. Consequently, DO has attracted interest in the areas of data integration and annotation, disease mapping, and computational analysis of disease associations[44–47]. Although DO embarks on encapsulating a comprehensive theory of disease[48], it still largely relies on non-mechanistic terms based on descriptive symptoms and unknown causes. Besides DO, other broad or focused harmonisation initiatives run in parallel, i.e., DisMaNET[31], Mondo[4,32], and SNOMED CT[4]. All these platforms aim to overcome the heterogeneity concerning disease terminology by providing a semantic model for integrating multiple ontologies. However, none of these efforts is fully comprehensive or sufficient to address the distinct needs of different stakeholders from academia, clinical centres, and industry. In fact, the existing classifications of diseases are not only agnostic about the underlying causal molecular pathology, but they also exacerbate the problem using generic and different labels, focusing in an alternating manner on organs, histology, time-course, and disease triggers.

Unlike the HUGO Gene Nomenclature Committee (HGNC), which set widely accepted and adopted guidelines for human genes[49], neither a consensus nor a nomenclature committee for disease classification exists so far.

Three inconsistency examples

To examine and visualise the dichotomy between different classifications together with the practical and innovation consequences, we focused on the following three common clinical phenotypes; asthma, stroke, and hypertension, and three resources; DO, MeSH, and ICD-10. In asthma, the current definition has diverged substantially from its historical origin. The term asthma derives from the ancient Greek ἅσθμα “laborious breathing”, an intrinsically symptom-based definition. Early clinical classification systems included allergic *vs.* non-allergic asthma, while recent classification has focused on endotypes (also known as mechanotypes, endophenotypes, subphenotypes or subtypes) based on the underlying immunological abnormalities, the impact of the inflammasome, and the epithelial barrier[50]. Although cellular and various multi-omic biomarkers for asthma have been proposed[51], they have neither been widely utilised in clinical research nor in care. In fact, for most patients, asthma treatment is based on the severity of symptoms[50]. Specifically, for mild to moderate asthmatics, bronchodilators (ß2 agonists) and/or inhaled corticosteroids are employed. Only in severe asthmatics, elements of disease subtyping are considered, e.g., with elevated

circulating eosinophils, anti-IL5 or anti-IL4 compounds[52]. To show the discrepancies between different classifications, asthma subtypes were extracted from the three mentioned resources as shown in (Fig. 2). In total, 23 clinical phenotypes of asthma were listed in the three resources, while only six jointly appear in two classifications; all others are singular, hence not matched to any of the other two classifications.

In hypertension, MeSH is more verbose including subtypes such as essential, malignant, pregnancy-induced, and renal among others (Supplementary Fig. 1). In contrast, MeSH is less descriptive of different taxonomic categorizations of ischaemic stroke (Supplementary Fig. 2). ICD, however, provides more granularity in terms of stroke classification considering cryptogenic, ischaemic, neonatal, perinatal, and postprocedural ischemia; whereas hypertension is not confined to hypertensive diseases (I10-I15) but is categorised under multiple disease classes (I, G, H, O, P, R, and Y). Therefore, diseases that are unified under the same tree branch in DO or MeSH are split up in ICD.

The limited overlap between these three resources warrants extreme caution when interpreting diagnoses across different clinical data sources, and mapping disease annotations to each other in research studies. While the highest-level definition (i.e., root node defining a major disease class) capturing shared symptoms (e.g., asthma, hypertension) are relatively coherent, the subsequent division of definitions varies greatly and inconsistently, highlighting the discord in each classification.

Biomedical research overlooks most clinical needs

The inconsistency perpetuates into research definitions and models of these three examples above; including (i) some clinical phenotypes not having correspondent preclinical disease models (intrinsic asthma; hypertension and comorbidities; vasospasm and stroke), and (ii) others overrepresented with a high bias (allergic asthma; renal hypertension; middle cerebral artery occlusion) (Fig. 2 and Supplementary Fig. 1 & 2). While in acute stroke, one can assume that the underlying causal mechanisms (although unknown) are most likely similar, it is unclear whether an animal model of asthma or hypertension also shares causal disease mechanisms and not just symptoms. Similarly, animal models for Alzheimer's and Parkinson's disease are supposedly mimicking human disease when the underlying molecular pathomechanisms of both indications have not been fully identified.

Clinicians have made several attempts to subclassify phenotypes such as asthma in a descriptive manner, ranging from different origins and triggers to different cell populations in the sputum of asthmatic patients. However, with respect to identifying molecular causes, geneticists often treat asthma as a single entity with which collaborative genome-wide associations are designed in order to obtain the large sample sizes required to detect genetic associations of small effect (the vast majority of complex disease genetic determinants). However, genetic differences likely contribute to phenotypic heterogeneity[53]. As a result, many primary disease-relevant genetic determinants of disease endophenotypes may have been missed.

Research traditions and unmet medical needs are highly uncoupled[5,54]. Obviously, geneticists, pre-clinical researchers and clinical staff are working in silos with hardly any cross-fertilisation, at least none towards clinically relevant results[5].

How big is the problem?

To further quantify the gap between these ontologies beyond the above mentioned three inconsistency examples, we fully analysed six widely used classification; ICD-10, MeSH, DO, OMIM, Orphanet, and Mondo. The mapping ratios for each pair of classifications were calculated and visualised in a heatmap (Fig. 3A and Supplementary Fig. 3). Specifically, each ratio is a normalised measure of how many unique terms from one ontology are

mapped to the corresponding terms in the second ontology (and thus, asymmetric). Moreover, most classifications suffer substantial information loss when mapping to one another. Translating OMIM to ICD-10 results in the loss of 54% of unique OMIM terms that have no mapping in ICD-10, while the opposite results in the loss of around 94% of ICD-10 terms. This scenario is indeed not surprising considering the purpose of both classifications, i.e., ICD-10 describes clinical observations, while OMIM covers phenotypes with genetic origins.

Out of 28,454 different disease terms in ICD-10, only 8% and 15% can be mapped to Orphanet and Mondo, respectively (Fig. 3B). Mondo, as an integrative ontology, fully maps to rare diseases in Orphanet. To further understand the implications of inconsistencies observed across different disease classifications, we describe various problematic mapping examples between ICD-10 and OMIM, as shown in Table 2 (Supplementary Table 2 for extended information). For instance, heart failure exemplifies disease terms that are in ICD-10, while not listed in OMIM. Another cross-mapping limitation arises from the fact that some diseases are named differently in both classifications. ICD-10 tends to overuse collective terms such as “other” or “unspecified” which hinders the homogenisation process amongst disease ontologies. Additionally, many drug-related disorders in OMIM do not exist in the ICD-10. Similarly, several OMIM diseases use the terms “progression of”, “protection against”, or “susceptibility to”, while no such terminology is used in ICD-10. This can result in wrong/imprecise mapping, or even a complete loss of data. Surprisingly, using pre-existing mapping sources, i.e., Mondo, KEGG, MalaCards, or Orphanet, has proven to be sometimes wrong, imprecise, or too general.

Thus, mapping one disease classification to another — a task frequently encountered in today’s data-driven medicine[55] and essential between biomedical fields using different ontologies — is by no means trivial often leading to inconsistencies in the interpretation of the data. Proper disease coding is crucial for processes where data need to be entered into and shared between clinical registries, genotype-phenotype databases, and biobanks[4]. In the realm of data sharing and management, the adoption of FAIR principles[56] becomes essential, while all aforementioned disease mapping-related and interoperability issues present a clear challenge.

Impact of the limitations in classifying diseases on medical innovation

The innovation roadblock is closely coupled to our current disease definitions, some of which have remained almost intact since the 19th century. In lieu of mechanism-based targets, drug discovery needs to stop assuming that by modulating a symptom the ultimate therapeutic goal will be reached, i.e., reducing glucose levels in blood to prevent diabetic complications. Although such approaches could maximally lead to symptomatic relief, the exact underlying pathomechanisms of the diseases are not yet elucidated. Even if a consensus would be reached between all current NCD nomenclatures, it would still be descriptive and none, apart from monogenic diseases, would allow for a mechanism-based definition and precision intervention. This is best evidenced by the imprecision of current medications, the most common interventions in health care.

Rare diseases are a role model to guide us where all other disease definitions need to move to. In most rare diseases, a precise gene variant has been identified providing at least a precise diagnosis and uniform ontology across all fields; clinical practice, genetics, and biomedical research. With the advent of gene editing, such mechanism-based precise diagnosis then allows increasingly for curative therapy, as opposed to chronically treating symptoms as we currently do. Of note, most monogenic diseases have symptoms in more than one organ, which is another argument supporting that an organ-based definition for more complex, oligogenic diseases makes little sense.

Thus, at least three knowledge gaps and roadblocks for innovation exist in our understanding of diseases:

- Organ/symptom- rather than mechanism-based definitions,
- Lack of relevance of pre-clinical disease models,
- Unsuccessful translation of pre-clinical research including genetics into patient benefit.

Systems medicine, the interactome and network pharmacology

Acknowledging that our current medical taxonomy is not sustainable, what is the alternative? We here propose to rapidly redefine our entire concept of disease by a purely mechanism-based approach towards identifying a molecular pathology. In the context of clinical trial design, there are calls for reforming human disease taxonomy, incorporating scientific advances in molecular and genetic medicine, and moving towards disease definitions based on molecular mechanisms[57,58]. The road towards such a mechanism-based disease ontology will require a systemic view of the human body (systems medicine), using disease-associated genes[59] as seeds (initial starting points) within the protein-protein interactome (PPI) to construct *de novo* signalling modules. Should these signalling modules be affected in a patient in a disease-relevant manner, we predict that in most cases lifestyle intervention will be sufficient to prevent the manifestation of a disease phenotype. However, in other cases, preventive or therapeutic pharmacological interventions will of course be still necessary.

Endophenotyping will dissect the current broad disease descriptions into specific PPI modules that are perturbed in certain individuals. Since these modules are small networks, they will be best treated with more than one drug in combination, i.e., taking advantage of network pharmacology[60]. For some endotypes where genes may not have been discovered, the molecular mechanisms shall be derived from triangulation through the genes or drugs of associated comorbidities., a process which is called multiscale modelling and network medicine.[60–63] Eventually all ‘common’, ‘complex’ or ‘chronic’ diseases and ‘syndromes’ will be endotyped and replaced by more precise mechanistic disease definitions[64].

Taking again asthma as an example of how the new classification would materialise, a fine-grained definition, for example referring to a group of patients with genetic risk factors and who have been smokers during a certain period of their lives, would be “chronic obstructive asthma characterised by (i) over-expansion and/or loss of (protein name), (ii) acquired/inherited mutations on (certain) interleukins, or (iii) high (based on a predefined threshold) inflammasome activity”. This definition would cover the clinical observations, the underlying genetic background of the patient, and the activity of genes within the cellular network such as the inflammasome. In the end, we will need to arrive at mechanistically defined endotypes which, depending on lifestyle and exposome, result in asthma-like symptoms, but most likely also other comorbidities. Each of these will be treated differently, targeting the causal mechanism or ideally by prevention.

As a proof of principle, in hypertension[65], we identified a disease module linked also to a higher risk in stroke[66], which is present in about one fourth of all patients with hypertension and can be fairly easily detected with a blood test. Functionally, it most likely relates to endothelium-dependent vasodilation (nitric oxide deficiency) and microvascular perfusion. In ischaemic stroke, the same module is affected, but in a different manner (nitric oxide overproduction). In this case, the mechanism can be addressed by network pharmacology[14] and appears to be uniform in all patients but confined to a post-stroke time-window. These findings are now in clinical phase II proof-of-concept trials within the European project REPO-TRIAL (repo-trial.eu) and platform REPO4EU (repo4.eu).

Thus, the way towards a molecular pathology of disease builds on genetic, omics and other big data approaches to develop a multi-scale network of all human diseases[63]. This is no longer manageable by a single specialist but requires network bioinformatics to analyse the clustering of different phenotypes due to shared genes, symptoms, drugs, or comorbidity associations[61]. These mechanistic definitions can then be captured in a molecular disease classification system[67]. For most NCDs, this will mean endo- or subtyping of the current organ- or symptom-based umbrella terms[65]. Most likely, such refined terms will be merged with identical molecular endotypes from other phenotypic umbrella terms[60]. In the end, there will be only molecular disease definitions annotated with several possible symptoms in different organs. At a clinical level, this will provide unprecedented molecular diagnostic and precision therapeutic opportunities with a substantially greater chance of achieving individual treatment success or even cure. To achieve this goal of a molecular pathology, data integration from all medical fields is key, even if the result would only be a transitory step towards a completely novel ontology. In the transition period, biomedical scientists and geneticists need to abandon their current independent terminologies in favour of the clinically used ICD. Only then, basic, genetic, and clinical data would be directly compatible.

Implications for clinical practice, research, and drug discovery

How will a conversion to mechanism-based classification affect clinical practice, patient benefit, biomedical research, and drug discovery? Systems-based or organ-agnostic approaches to disease have already started driving innovation in oncology, although, with the exception of haematologic and lymphoid malignancies[68], the true leap forward is still missing. To modern clinicians, it is obvious that a tumour is better defined and treated based on its individual mutagenic burden, rather than the primary organ or histology[69]. Similar insights have entered into primary immune diseases[70], as well as comorbid cardiometabolic and neurological phenotypes[60,61]. While the benefit for patients to move from chronic symptomatic therapy to cure or ideally to early detection and prevention is obvious, the outlooks for clinicians are more profound. Organ-based specialisation will have to be left behind, and integrative, general medicine approaches will be needed. This requires of course substantial support from decision algorithms and support systems such as Symptoma (symptoma.com) or Ada Health (ada.com) among others.

Medical research will need to move more to the human-centred world and eventually from disease to prevention. Pharmaceutical research will at first be helped by the new precise molecular disease definitions and diagnostics for patient inclusion, which de-risk drug discovery considerably[71,72]. The earliest example and application of a mechanistic redefinition of disease and mechanistic drug approval are pembrolizumab (Keytruda®) and larotrectinib (Vitrakvi®), two cancer drugs based on gene variants rather than an organ-based or histological tumour definition[73,74].

Individual markets will be smaller as compared to the current umbrella term indications and for many molecular indications, no new drugs may be necessary and existing ones may be repurposed. The latter is an approach that both the FDA and EMA support and closely monitor. We may then in the not too far future reach the point where we have in principle all the drugs we need. Why should we need to eternally look for new drugs, anyway?

Conclusions

To enter the era of true precision medicine and prevention, our current wild growth of disease definitions is the biggest roadblock. Medicine urgently needs an international Consensus Nomenclature Committee for disease classification. This may start off with ICD to ensure buy-in from as many clinicians as possible. Evidence that complex disease phenotypes will benefit from molecular definitions and subtypes comes from the 7,000 rare diseases most of which are named after the responsible gene or protein[75], providing both a precise

diagnosis and options for a curative strategy[76]. To reduce expert mapping workload, eventually, automatic ontology curation with similarity-based reasoning may be employed. To map this to molecular data, geneticists, biomedical and clinical researchers need to associate their data to as many ontologies as possible with appropriate cross-references, even if they still include organ- and symptom-based descriptions.[77,78] The final outcome, a mechanism-based reclassification of non-communicable diseases, will define patients based on molecular endotypes leading to different phenotypes (currently, our disease definitions; then, just symptoms), which can be diagnosed by genetic or precision diagnostic tools and treated by precision, network pharmacology, or ideally precision prevention through early and targeted lifestyle intervention.

Acknowledgements

The REPO-TRIAL project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 777111; the REPO4EU project, from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101057619. This reflects only the authors' views and the European Commission is not responsible for any use that may be made of the information it contains (H.H.H.S., A.I.C, E.G). A.I.C. was also supported by the Förderprogramm der Corona-Stiftung im Stifterverband and the BMBF funding (ref. 01EJ2205D). Z.M. is funded by a full scholarship [No 40463/2019] from the Ministry of higher education of the Arab Republic of Egypt.

Tables

Table 1. Annotation and classification systems used in clinical practice and biomedical research*.

Dataset name	Full name	Number of disease terms	Ontology structure/curation	Application field
SNOMED CT	Systematized Nomenclature of Medicine - Clinical Terms	78,213	Tree based, manual curation	Electronic health records Natural language processing Medical records systems, Computerized forms and records control
Mondo	Monarch Disease Ontology	43,804	Graph based, semi-automatic curation	Computational biology Whole genome sequencing Translational medical research
DO	Disease Ontology	17,667	Graph based, manual curation	Computational biology Molecular sequence annotation Gene expression profiling
Orphanet	Orphanet Rare Diseases Ontology	15,205	Graph based, manual curation	Rare diseases Orphan drug production
MeSH	Medical Subject Headings	12,750	Tree based, manual curation	Information storage and retrieval Abstracting and indexing Natural language processing
ICD10 ICD11	International Classification of Diseases	12,542 36,275	Tree based, manual curation	Electronic health records Clinical coding Qualitative research
OMIM	Online Mendelian Inheritance in Man	9,575	Flat list, manual curation	Computational biology Genomics Genome-wide association study

* This table provides a comparison between different classification systems including; Systematized Nomenclature of Medicine – Clinical Terms (SNOMED CT), The Mondo disease ontology (Mondo), Human Phenotype Ontology (HPO), Disease Ontology (DO), Orphanet, Medical Subject Headings (MeSH), International Classification of Diseases-10/11 (ICD10/11), and Online Mendelian Inheritance in Man (OMIM for genetic phenotypes). As of 2022-02-08.

406 **Table 2. Type of problems when mapping ICD10 and OMIM disease terms*.**

Mapping problem	ICD10	OMIM
Different related diseases are linked to different OMIM codes while only one is used in ICD-10	Congenital cataract (Q12.0)	Cataract, multiple types (115650, 611391, 116800, 115665, 116200, 115700, 116100, 605749, 613763, 601885, 611544, 605387, 615277, 600881, 610202, 611597) Cataract, autosomal recessive (614691,610019)
Several rare diseases and syndromes are present in OMIM but are missing in ICD-10	Does not exist in ICD-10	Hyperproinsulinemia (616214)
Several diseases in ICD10 do not exist in OMIM	Heart failure (I50)	Does not exist in OMIM
Several entries in OMIM have different descriptions than those found in ICD-10	Aneurysm of heart (I25.3)	Aneurysm of interventricular septum (105805)
Indications present in ICD-10 are not specified and frequently use collective terms, while OMIM entries are highly specific	Other specified diabetes mellitus (E13)	Diabetes mellitus, insulin-resistant, with acanthosis nigricans (610549) Rabson-Mendenhall syndrome (262190)

Various diseases are specified by the gene/locus in OMIM and cannot be mapped to specific codes in ICD-10	In ICD-10, acute myocardial infarction (I21) is classified based on the location Acute transmural myocardial infarction of anterior wall (I21.0) Acute transmural myocardial infarction of inferior wall (I21.1) Acute transmural myocardial infarction of other sites (I21.2)	Myocardial infarction, susceptibility to (608446, 608557)
Diseases classified in OMIM based on therapy administered and are not included in ICD-10	Coronary vasodilators, not elsewhere classified (Y52.3)	Sublingual nitroglycerin, susceptibility to poor response to (100650)
Diseases specified in OMIM using “progression of”, “protection against”, and “susceptibility to” are only mentioned by disease name in ICD-10	Other conduction disorders (I45)	Cardiac conduction defect, susceptibility to (115080)
Some classifications in OMIM are confusing and/or can receive two different codes in ICD-10	“Alzheimer disease” is found under the code G30 (Chapter 6 Diseases of the nervous system) while “Dementia in Alzheimer disease” is present under F00 (Chapter 5 V Mental and behavioural disorders)	Alzheimer disease (607822, 104300, 104310, 606889, 608907, 615590, 602096, 605526, 606187, 607116, 609636, 609790, 611073, 615080, 615711, 300756, 605055, 611152, 611154, 611155, 604154, 607413)
Pre-existing mapping is sometimes wrong, imprecise, or too general	Nephrotic syndrome, type 7 (615008) is mapped to N00.5 (Does not exist in ICD-10) in Mondo. MalaCards and Orphanet both map it to N00.5 and D58.8 (Other specified hereditary haemolytic anaemias). While it is mapped to N0.4 (Nephrotic syndrome) in KEGG which is too general.	Nephrotic syndrome, type 7 (615008)

* This table provides representative examples of ten of the main problems encountered while mapping disease terms between two disease classification systems: International

Classification of Diseases 10 (ICD-10) and Online Mendelian Inheritance in Man (OMIM). (For an extended version, see Supplementary Table 2).

Figures and figure legends

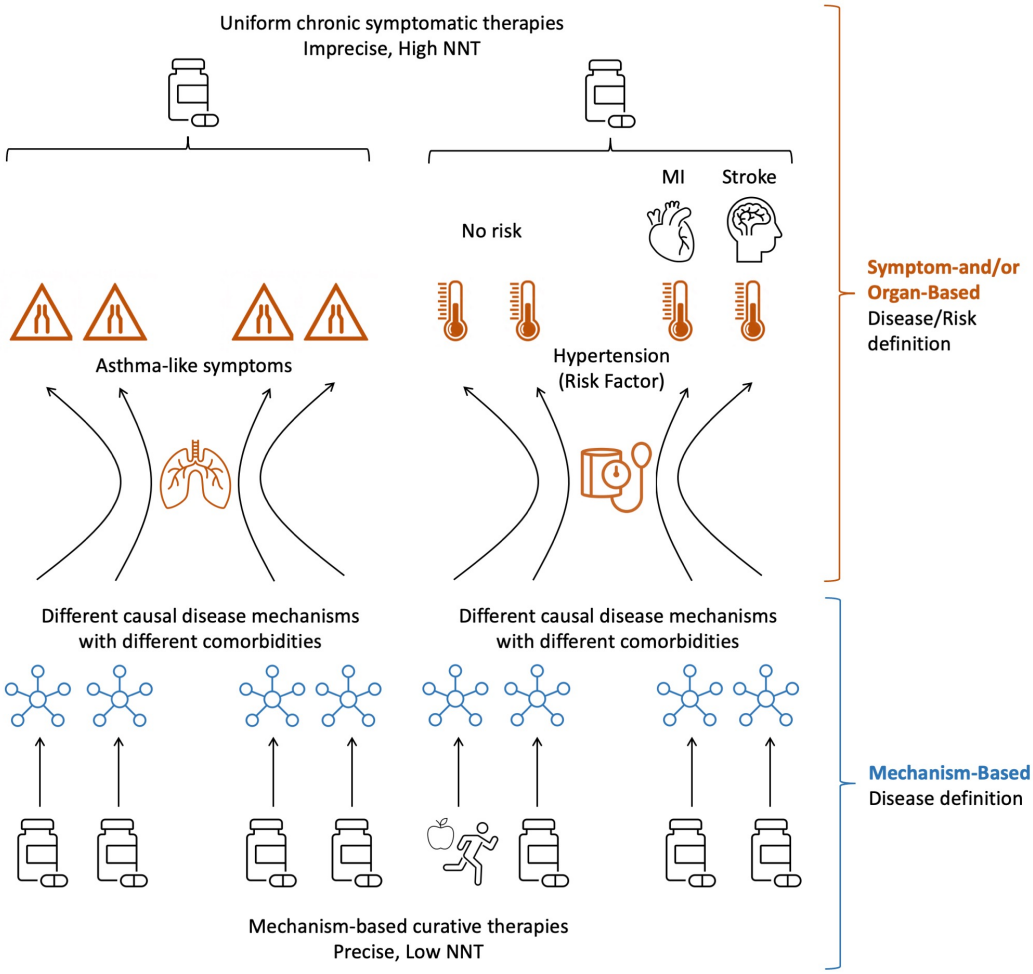

Fig. 1: Organ/symptom-based versus mechanism-based disease definitions and their consequences for therapeutic precision.
In orange, the current symptom- and/or organ-based approach to define diseases (left asthma-like symptoms, right hypertension as risk factor) leading to imprecise therapies with high numbers needed to treat (NNT) asthma or prevent the consequences of hypertension; myocardial infarction (MI) or stroke. Note that many patients with hypertension are not at risk but are treated as well.
In blue, mechanism-based approaches[79], where different mechanisms can lead to similar symptoms, but are treated differently, i.e., in a mechanism-based manner offering a higher degree of precision and possibility to cure[60]

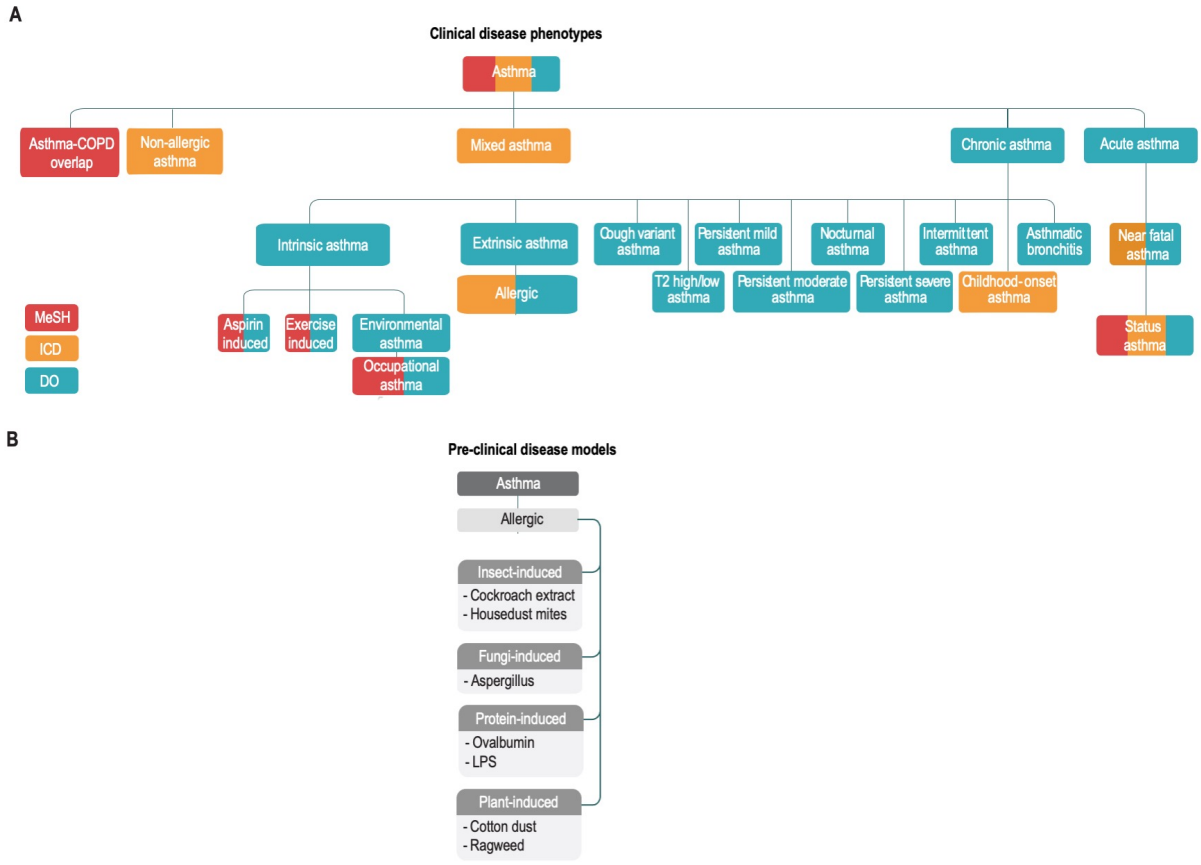

Fig. 2: Mismatch of clinical phenotypes and preclinical models of asthma.
(A) The results for the keyword "asthma" when searched in three popular disease classifications; Disease Ontology (DO, blue), Medical Subject Headings (MeSH, red) and the International Classification of Diseases 10 (ICD-10, orange). (B) Examples of preclinical disease models of asthma and their limited overlap with human phenotypes of asthma.

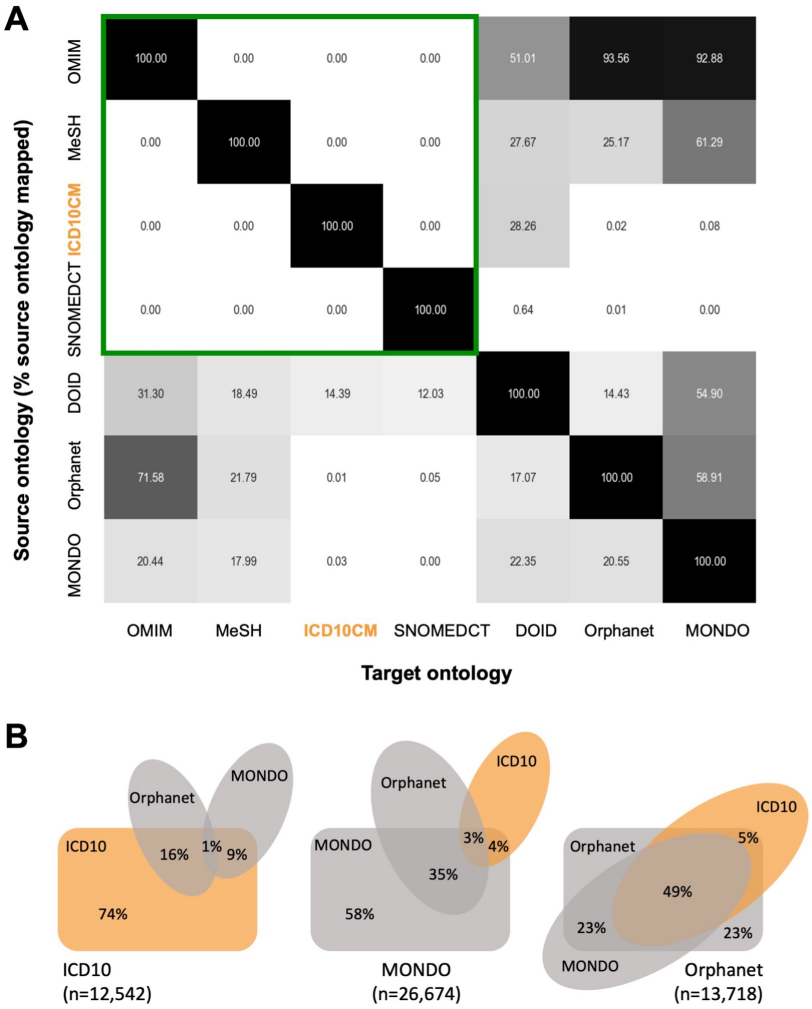

Fig. 3: Mapping coverage between different disease ontologies.
(A) on the heat map, the labels indicate the percentage of the pairwise ontology mapping coverage of the source ontology (y-axis) relative to the target ontology (x-axis). The percentages are not symmetrical due to differences in the total number of terms of each ontology, and the fact that mappings are often not one-to-one. The green bounding box highlights zero coverage with direct mappings between the ontologies (distance = 1), while indirect mapping provides non-zero coverage (for more details about direct and indirect mappings, see supplementary Fig. S3). (B) three-way mapping coverage between the ICD10, Mondo, and Orphanet ontologies. There is one Venn diagram for each base ontology (rectangles) and the overlap with the other two ontologies (ovals). The total number of terms in the base ontology is displayed below the corresponding Venn diagram. The percentages shown on the Venn diagrams represent the coverage relative to the base ontology, intersections between the ontologies, and the remainder of the unmapped percentage of the base ontology.

References

1. Organization, W. H. & Others. Noncommunicable diseases progress monitor 2022. (2022).
2. Schork, N. J. Personalized medicine: Time for one-person trials. *Nature* **520**, 609–611 (2015).
3. Goh, K.-I. *et al.* The human disease network. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 8685–8690 (2007).
4. Sollie, A. *et al.* A New Coding System for Metabolic Disorders Demonstrates Gaps in the International Disease Classifications ICD-10 and SNOMED-CT, Which Can Be Barriers to Genotype--Phenotype Data Sharing. *Hum. Mutat.* **34**, 967–973 (2013).
5. Contopoulos-Ioannidis, D. G., Ntzani, E. E. & Ioannidis, J. P. A. Translation of highly promising basic science research into clinical applications. *Am. J. Med.* **114**, 477–484 (2003).
6. Armstrong, G. L. Trends in Infectious Disease Mortality in the United States During the 20th Century. *JAMA* vol. 281 61 Preprint at https://doi.org/10.1001/jama.281.1.61 (1999).
7. Woolf, S. H. & Schoomaker, H. Life Expectancy and Mortality Rates in the United States, 1959-2017. *JAMA* **322**, 1996–2016 (2019).
8. Hiam, L., Dorling, D. & McKee, M. Things Fall Apart: the British Health Crisis 2010-2020. *Br. Med. Bull.* **133**, 4–15 (2020).
9. Schmidt, H. H. H. *The End of Medicine as We Know it -- and why Your Health Has a Future*. (Springer Nature, 2022).
10. Wieseler, B., McGauran, N. & Kaiser, T. New drugs: where did we go wrong and what can we do better? *BMJ* l4340 Preprint at https://doi.org/10.1136/bmj.l4340 (2019).
11. Echt, D. S. *et al.* Mortality and Morbidity in Patients Receiving Encainide, Flecainide, or Placebo. *New England Journal of Medicine* vol. 324 781–788 Preprint at https://doi.org/10.1056/nejm199103213241201 (1991).
12. Major cardiovascular events in hypertensive patients randomized to doxazosin vs chlorthalidone: the antihypertensive and lipid-lowering treatment to prevent heart attack trial (ALLHAT). ALLHAT Collaborative Research Group. *JAMA* **283**, 1967–1975 (2000).
13. Scannell, J. W., Blanckley, A., Boldon, H. & Warrington, B. Diagnosing the decline in pharmaceutical R&D efficiency. *Nat. Rev. Drug Discov.* **11**, 191–200 (2012).
14. Casas, A. I. *et al.* From single drug targets to synergistic network pharmacology in ischemic stroke. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 7129–7136 (2019).
15. Freedman, L. P., Cockburn, I. M. & Simcoe, T. S. The Economics of Reproducibility in Preclinical Research. *PLoS Biol.* **13**, e1002165 (2015).
16. Brembs, B. Prestigious Science Journals Struggle to Reach Even Average Reliability. *Front. Hum. Neurosci.* **12**, 1–7 (2018).
17. Noorden, V. Science publishing: The trouble with retractions. *Nature Publishing Group UK* http://dx.doi.org/10.1038/478026a (2011) doi:10.1038/478026a.
18. Agache, I., Akdis, C., Jutel, M. & Virchow, J. C. Untangling asthma phenotypes and endotypes. *Allergy* **67**, 835–846 (2012).
19. Tomassen, P. *et al.* Inflammatory endotypes of chronic rhinosinusitis based on cluster analysis of biomarkers. *J. Allergy Clin. Immunol.* **137**, 1449–1456.e4 (2016).
20. Ghiassian, S. D. *et al.* Endophenotype Network Models: Common Core of Complex Diseases. *Sci. Rep.* **6**, 27414 (2016).
21. Hinchcliff, M. & Mahoney, J. M. Towards a new classification of systemic sclerosis. *Nature reviews. Rheumatology* vol. 15 456–457 (2019).
22. Lüscher, T. F. Lumpers and splitters: The bumpy road to precision medicine. *Eur. Heart J.* **40**, 3292–3296 (2019).

23. Van Rees, R. Clarity in the usage of the terms ontology, taxonomy and classification. *Civil Engineering -New York then Reston-* **20**, (2008).
24. Israel, R. A. The history of the International Classification of Diseases. *Health Bull (Edinb)* **49**, 62–66 (1991).
25. McKusick, V. A. Mendelian Inheritance in Man and its online version, OMIM. *Am. J. Hum. Genet.* **80**, 588–604 (2007).
26. Rogers, F. B. Medical subject headings. *Bull. Med. Libr. Assoc.* **51**, 114–116 (1963).
27. Arguello Casteleiro, M. *et al.* Deep learning meets ontologies: experiments to anchor the cardiovascular disease ontology in the biomedical literature. *J. Biomed. Semantics* **9**, 13 (2018).
28. Aymé, S. & Schmidtke, J. Networking for rare diseases: a necessity for Europe. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz* **50**, 1477–1483 (2007).
29. Schriml, L. M. *et al.* Human Disease Ontology 2018 update: classification, content and workflow expansion. *Nucleic Acids Res.* **47**, D955–D962 (2019).
30. Smith, B. *et al.* The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. *Nat. Biotechnol.* **25**, 1251–1255 (2007).
31. Jallon, J. M., Risler, Y. & Iwatsubo, M. Beef liver L-Glutamate dehydrogenase mechanism: presteady state study of the catalytic reduction of 2.oxoglutarate by NADPH. *Biochem. Biophys. Res. Commun.* **67**, 1527–1536 (1975).
32. Vasilevsky, N. A. *et al.* Mondo: Unifying diseases for the world, by the world. *bioRxiv* (2022) doi:10.1101/2022.04.13.22273750.
33. Malone, J. *et al.* Modeling sample variables with an Experimental Factor Ontology. *Bioinformatics* **26**, 1112–1118 (2010).
34. Drösler, S. E., Romano, P. S., Tancredi, D. J. & Klazinga, N. S. International comparability of patient safety indicators in 15 OECD member countries: a methodological approach of adjustment by secondary diagnoses. *Health Serv. Res.* **47**, 275–292 (2012).
35. Khokhar, B. *et al.* Systematic review of validated case definitions for diabetes in ICD-9-coded and ICD-10-coded data in adult populations. *BMJ Open* **6**, e009952 (2016).
36. McCormick, N., Bhole, V., Lacaille, D. & Avina-Zubieta, J. A. Validity of Diagnostic Codes for Acute Stroke in Administrative Databases: A Systematic Review. *PLoS One* **10**, e0135834 (2015).
37. Peng, M. *et al.* Coding reliability and agreement of International Classification of Disease, 10th revision (ICD-10) codes in emergency department data. *Int J Popul Data Sci* **3**, 445 (2018).
38. Xiao, A. Y. *et al.* The Use of International Classification of Diseases Codes to Identify Patients with Pancreatitis: A Systematic Review and Meta-analysis of Diagnostic Accuracy Studies. *Clin. Transl. Gastroenterol.* **9**, 191 (2018).
39. Schwarze, K., Buchanan, J., Taylor, J. C. & Wordsworth, S. Are whole-exome and whole-genome sequencing approaches cost-effective? A systematic review of the literature. *Genet. Med.* **20**, 1122–1130 (2018).
40. Zhou, X., Menche, J., Barabási, A.-L. & Sharma, A. Human symptoms-disease network. *Nat. Commun.* **5**, 4212 (2014).
41. Kim, S., Yeganova, L. & Wilbur, W. J. Meshable: searching PubMed abstracts by utilizing MeSH and MeSH-derived topical terms. *Bioinformatics* **32**, 3044–3046 (2016).
42. Huang, M., Névéol, A. & Lu, Z. Recommending MeSH terms for annotating biomedical articles. *J. Am. Med. Inform. Assoc.* **18**, 660–667 (2011).
43. Funk, M. E. & Reid, C. A. Indexing consistency in MEDLINE. *Bull. Med. Libr. Assoc.* **71**, 176–183 (1983).
44. Croft, D. *et al.* The Reactome pathway knowledgebase. *Nucleic Acids Res.* **42**, D472–7 (2014).

45. Köhler, S. *et al.* The Human Phenotype Ontology project: linking molecular biology and disease through phenotype data. *Nucleic Acids Res.* **42**, D966–74 (2014).
46. Žitnik, M., Janjić, V., Larminie, C., Zupan, B. & Pržulj, N. Discovering disease-disease associations by fusing systems-level molecular data. *Sci. Rep.* **3**, 3202 (2013).
47. Cheng, L. *et al.* DisSim: an online system for exploring significant similar diseases and exhibiting potential therapeutic drugs. *Sci. Rep.* **6**, 30024 (2016).
48. Schriml, L. M. et al. The Human Disease Ontology 2022 update. Nucleic Acids Res. 50, D1255–D1261 (2022).
49. Bruford, E. A. *et al.* Guidelines for human gene nomenclature. *Nat. Genet.* **52**, 754–758 (2020).
50. Agache, I. & Akdis, C. A. Precision medicine and phenotypes, endotypes, genotypes, regiotypes, and theratypes of allergic diseases. *J. Clin. Invest.* **129**, 1493–1503 (2019).
51. Custovic, A., Henderson, J. & Simpson, A. Does understanding endotypes translate to better asthma management options for all? *J. Allergy Clin. Immunol.* **144**, 25–33 (2019).
52. Colas, L., Hassoun, D. & Magnan, A. Needs for Systems Approaches to Better Treat Individuals With Severe Asthma: Predicting Phenotypes and Responses to Treatments. *Front. Med.* **7**, 98 (2020).
53. Ferreira, M. A. R. *et al.* Genetic Architectures of Childhood- and Adult-Onset Asthma Are Partly Distinct. *Am. J. Hum. Genet.* **104**, 665–684 (2019).
54. Ioannidis, J. P. A. Why Most Clinical Research Is Not Useful. *PLoS Med.* **13**, e1002049 (2016).
55. Lehne, M., Sass, J., Essenwanger, A., Schepers, J. & Thun, S. Why digital medicine depends on interoperability. *NPJ Digit Med* **2**, 79 (2019).
56. Wilkinson, M. D. *et al.* Addendum: The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* **6**, 6 (2019).
57. Kola, I. & Bell, J. A call to reform the taxonomy of human disease. *Nat. Rev. Drug Discov.* **10**, 641–642 (2011).
58. Hofmann-Apitius, M., Alarcón-Riquelme, M. E., Chamberlain, C. & McHale, D. Towards the taxonomy of human disease. *Nat. Rev. Drug Discov.* **14**, 75–76 (2015).
59. Menche, J. *et al.* Disease networks. Uncovering disease-disease relationships through the incomplete interactome. *Science* **347**, 1257601 (2015).
60. Nogales, C. *et al.* Network pharmacology: curing causal mechanisms instead of treating symptoms. *Trends Pharmacol. Sci.* (2021) doi:10.1016/j.tips.2021.11.004.
61. Langhauser, F. *et al.* A diseasome cluster-based drug repurposing of soluble guanylate cyclase activators from smooth muscle relaxation to direct neuroprotection. *NPJ Syst Biol Appl* **4**, 8 (2018).
62. Barabási, A.-L., Gulbahce, N. & Loscalzo, J. L. Network medicine: a network-based approach to human disease. *Nat. Rev. Genet.* **12**, 56–68 (2011).
63. Sadegh, S. *et al.* Network medicine for disease module identification and drug repurposing with the NeDRex platform. *Nat. Commun.* **12**, 6848 (2021).
64. Vallance, P. An audience with Patrick Vallance. *Nat. Rev. Drug Discov.* **9**, 834 (2010).
65. Elbatreek, M. H. *et al.* NOX5-induced uncoupling of endothelial NO synthase is a causal mechanism and theragnostic target of an age-related hypertension endotype. *PLoS Biol.* **18**, e3000885 (2020).
66. Casas, A. I. *et al.* Calcium-dependent blood-brain barrier breakdown by NOX5 limits postreperfusion benefit in stroke. *J. Clin. Invest.* (2019) doi:10.1172/JCI124283.
67. Prieto Santamaría, L. *et al.* Classifying diseases by using biological features to identify potential nosological models. *Sci. Rep.* **11**, 21096 (2021).
68. Elmore, L. W. *et al.* Blueprint for cancer research: Critical gaps and opportunities. *CA Cancer J. Clin.* **71**, 107–139 (2021).

69. Sanchez-Vega, F. *et al.* Oncogenic Signaling Pathways in The Cancer Genome Atlas. *Cell* **173**, 321–337.e10 (2018).
70. Ellinghaus, D. *et al.* Analysis of five chronic inflammatory diseases identifies 27 new associations and highlights disease-specific patterns at shared loci. *Nat. Genet.* **48**, 510–518 (2016).
71. Pushpakom, S. et al. Drug repurposing: progress, challenges and recommendations. Nat. Rev. Drug Discov. 18, 41–58 (2018).
72. Ashburn, T. T. & Thor, K. B. Drug repositioning: identifying and developing new uses for existing drugs. Nat. Rev. Drug Discov. 3, 673–683 (2004).
73. Sweis, R. F. & Luke, J. J. Mechanistic and pharmacologic insights on immune checkpoint inhibitors. Pharmacol. Res. 120, 1–9 (2017).
74. Pollack, M. et al. Transforming approaches to treating TRK fusion cancer: historical comparison of larotrectinib and histology-specific therapies. Curr. Med. Res. Opin. 37, 59–70 (2021).
75. Buphamalai, P., Kokotovic, T., Nagy, V. & Menche, J. Network analysis reveals rare disease signatures across multiple levels of biological organization. Nat. Commun. 12, 6306 (2021).
76. Braga, L. A. M., Conte Filho, C. G. & Mota, F. B. Future of genetic therapies for rare genetic diseases: what to expect for the next 15 years? Therapeutic Advances in Rare Disease 3, 26330040221100840 (2022).
77. Zhou, H., Wang, Y. & Cao, K. Fuzzy D-S Theory Based Fuzzy Ontology Context Modeling and Similarity Based Reasoning. 2013 Ninth International Conference on Computational Intelligence and Security Preprint at https://doi.org/10.1109/cis.2013.154 (2013).
78. Gan, M., Dou, X. & Jiang, R. From Ontology to Semantic Similarity: Calculation of Ontology-Based Semantic Similarity. The Scientific World Journal vol. 2013 1–11 Preprint at https://doi.org/10.1155/2013/793091 (2013).
79. Schmidt, H. H. H. W. & Menche, J. The regulatory network architecture of cardiometabolic diseases. Nat. Genet. 54, 2–3 (2022).

Author contributions

E.G., and H.H.H.W.S. conceptualised and finalised this manuscript; Z.M.M., M.E., A.I.C., C.N., M.P.P, A.P, and Z.P., conducted disease mapping; E.A., and A.A.H. analysed ontologies; H.H.H.W.S., A.I.C., Z.M.M., E.A., A.A.H., and M.E. prepared figures or tables; Z.M.M., A.R.G., J.B., A.W., and E.K.S. edited the manuscript.

Competing interests

The authors declare no competing interests.