

COPD diagnosis: genome first

Lawrence Grouse

Department of Neurology, University of Washington School of Medicine, Washington, USA

Correspondence to: Lawrence Grouse, MD, PhD. Executive Director of the International COPD Coalition, Department of Neurology, University of Washington School of Medicine, 1959 NE Pacific Ave., Rm RR650, Box 356465, Seattle, Washington 98195-6465, USA. Email: lgrouse@uw.edu.

Submitted Sep 26, 2015. Accepted for publication Sep 26, 2015.

doi: 10.3978/j.issn.2072-1439.2015.10.38

View this article at: <http://dx.doi.org/10.3978/j.issn.2072-1439.2015.10.38>

The US Institute of Medicine has just released a Report documenting the problems in the US relating to the frequency with which physicians make incorrect diagnoses of their patients' illnesses (1). The Report included the calculation that on average all Americans could expect to receive an incorrect diagnosis from a physician at least once in their life. It explained that this error could be of minor concern or it could lead to unnecessary surgery or a lifetime of expensive, unneeded medical care. Patients who receive this information are very concerned about what misdiagnosis they may have received.

Having experienced the US medical care system, I consider this estimate by the Institute of Medicine to be a credible, even a conservative one. But they were considering errors in differential diagnosis of diseases such as acute MI and GI reflux, which could be mistaken for each other. What about diseases that share an overlap with another disease like the Asthma COPD Overlap Syndrome (ACOS) (2)? There is not even a diagnostic algorithm that is agreed on for making this diagnosis.

Another problem in making a correct diagnosis is that some conditions we call "diseases" are not actually single, unitary diseases at all. They are "wastebasket" diagnoses that contain many different sub-types of diseases, often with different pathophysiologies, all under the same name. Increasing evidence leads to the conclusion that COPD is just such a misleading diagnosis. Dr. Stephen Rennard and his colleagues recently published a study of more than 2,000 patients diagnosed as having COPD, and they found at least five subgroups of disease with different prognoses and requiring different treatments among the patients (3).

Diagnosing patients' diseases accurately is a primary responsibility of medical practice. If an incorrect diagnosis is made then the treatment the patient receives is likely to be

harmful. There should be a higher standard for diagnoses. The most valuable diagnoses will include both the causal pathophysiology and genetic basis of the disease.

The diagnosis of disease has historically been based on the observable characteristics or traits of the patient: their phenotype. The genetic cause of disease is difficult to analyze based only on patients' phenotypes. However, with the increasing access to patients' genotypes through genome sequencing, it may be possible to determine the genetic cause of a patient's disease, even in complex diseases that are not currently understood. This will provide an opportunity to identify specific therapy for the causative abnormality.

With the information gained from the sequencing of the human genome through The Human Genome Project and the vast improvement in DNA sequencing technology, physicians with access to their patients' genotype will be better able to diagnose their illnesses. Testing the genomes of patients with complex diseases through genomic microarray platforms and exomic and whole genome sequencing can reveal the genomic abnormalities that lead to these conditions (4). A clearer understanding of disease can be had through correlation between the genomic structure and the resulting patient phenotype(s). In many cases a genotype first approach to diagnosis that is now possible will improve diagnosis of complex diseases.

Complex disease is defined as a phenotype that is not caused by a single gene mutation but by many individual gene events, and with a significant contribution from environmental factors. Classical approaches to complex disease have identified patients with similar phenotypes and have attempted to identify the common causative mutation. In most cases the heritability of complex disease is unresolved using this strategy. COPD, for example, can be caused by abnormal alleles in cases of Alpha-1

Antitrypsin Deficiency and in other familial COPD cases, but its association with tobacco usage and other toxic inhalation is also clear in many patients. Epigenetic factors may play a role in the development of COPD (5). The genetic heterogeneity exemplified by COPD requires a shift in the approach to studying the genetics of the disease. By sequencing patients' exomes and comparing them with controls, a variety of genetic causes of the disease can often be identified. The exome is the portion of the genome that codes information for protein synthesis. Investigators can leverage technology to genetically classify subtypes of COPD. This involves candidate causal gene discovery and determination of pathogenicity, comprehensive clinical phenotyping, and resolution of genetic background effects.

The initial application of this genotype first approach to diagnosis is particularly applicable for complex diseases in which the molecular causes are not currently understood. This includes a substantial portion of patients' illnesses. Up to now, determining patients' genotypes has not been part of usual diagnostic practice. Take a typical patient with an infectious disease, for example. A young patient with an *S. pneumoniae* lung infection is diagnosed by a history and physical exam, laboratory and lung imaging information, sputum gram stain and bacterial cultures. The causative pathogen is identified and appropriate antibiotic therapy can cure the infection. However, in the future, genotyping of such a patient may well be important to explain, for example, why a young person is susceptible to bacterial infections. Do they have a genetically-related immune impairment? Perhaps in the future it will be important to examine the genomic risk factors even for common diseases that are currently being treated satisfactorily. With the rapid improvement of DNA sequencing and its diminishing cost, a copy of the sequence of each patient's genome may become a part of all medical charts.

Consider how many diseases physicians know or think that they know. How many of them are actually understood in terms of biological/genetic mechanisms of causation? For every disease like atherosclerotic coronary artery disease in which there is information about the causative intimal lesions that block the arteries and how they develop, there are other diseases like autism, lupus, and COPD, whose molecular mechanisms of causation are unknown and for which no curative therapy exists. There are countless diseases that have names and descriptions, which are similarly opaque to physicians who try to diagnose them, understand their etiologies, and treat them. Often the silos that medical specialties construct between disciplines

limit physicians' ability to understand the similarities and differences among diseases.

Although we are taught about Mendelian genetics, there is an ocean of other genetic pathology that is still outside our ken. We talk about COPD as if it were a single disease when it is clear that there are many subtypes of this syndrome. Physicians are still diagnosing COPD by detecting signs of airflow obstruction when a patient blows into a tube. This is not likely to provide molecular pathophysiologic information. We don't understand the molecular mechanisms of causation for any of the different subtypes of COPD.

On January 20, 2015, US President Barack Obama put forward his new initiative called the "Precision Medicine Initiative," which relies on genomics to "deliver the right treatment at the right time." He said, "I want the country that... mapped the human genome to lead a new era of medicine." He added that this kind of personalized medicine could lead to cures for cancer and diabetes. The importance of genomics will be the lynchpin in establishing sound scientific diagnoses that offer logical targets for therapy of causative molecular lesions.

Genomic mutations and rearrangements drive human evolution and phenotypic diversity as well as causing human disease so it is highly likely that this source of disease will always be with us. The human genome is not a garden of similar structural genes, it is a wild jungle of DNA sequences mutating, crossing over, and traveling around the genome with unpredictable results. The genome is in turmoil, casting the dice during cell division to select the final sequences. The resulting genome that survives in the existing environment is the winner. Genes are in contact with like genes throughout the genome through their similar sequences and affinities for certain protein structures. Huge sequences can be duplicated and inserted near other major duplications when they are summoned or they may be deleted or inverted. It is no surprise that individuals are each so different and also that so many human genetic abnormalities exist.

The impact of changes in genome structure is illustrated by their role in human genetic disease. The impact of Copy Number Variants (CNVs) is larger than single nucleotide variations in genome evolution. They have been implicated in a number of neurodevelopmental disorders such as autism spectrum disorders (ASD) (6). Segmental duplications are important sources of genomic instability. The contribution of *de novo* mutations of DNA to human disease and evolution is also important. Studying

genome-wide mutation rates and patterns is important for understanding mutation origins, locating hotspots for such events, estimating disease risk, and interpreting novel disease-associated mutations. Studies have convincingly shown that large and dramatic genome changes introduced by large structural mutations can be associated with a multitude of pathological conditions. Sequencing the exome of genes in patients with complex diseases such as autism and comparing them with exomes of control patients reveals numerous mutations among the patients. Many of these mutated patient genes are found to belong in networks of related genes, to be in physical contact with each other, and to be coexpressed. These are genes that have been found to be involved in the causation of ASD (7). This genotype first approach to diagnosis goes to the source of diagnostic information: the patient's DNA (8). While this approach will only be applicable to some complex diseases, it should be beneficial for many patients who are currently in need.

There is a recent example of the value of such a genetic analysis in the commonly occurring complex disease of obesity (9). A key gene mutation was seen in a regulatory gene for fat storage in patients with obesity. It was located in the FTO region of the genome, which harbors the strongest genetic association with human obesity. The mutation in the adipocytes led to fat storage rather than fat burning, which contributed to the obesity. The gene's code could be corrected in the adipocytes *in vitro* using the CRISPR-Cas9 editing system. Only by identifying the genetic lesion for the disorder could the repair be accomplished. Obviously, such repairs in actual patients will probably not be possible for many years, if at all. But unless we identify the genetic lesions that cause complex diseases, our ability to do developmental research and testing for safety and effectiveness of treatments will never be possible.

We have to admit that our approach to diagnosing what we call COPD leaves a great deal to be desired, although CT imaging appears promising. Our current treatments do not affect the basic pathologies or patient outcomes of COPD patients. Although we have to provide the best palliative care possible for COPD patients, advocates for COPD patients need to emphasize the importance of research to understand this heterogeneous mix of

pathologies we call COPD and diagnose their molecular and genetic causes so that meaningful therapy can become available.

Acknowledgements

The author would like to express his gratitude to Prof. Evan Eichler for his helpful information.

Footnote

Conflicts of Interest: The author has no conflicts of interest to declare.

References

1. Balogh EP, Miller BT, Ball JR. Improving Diagnosis in Health Care. Institute of Medicine, National Academies Press, 2015.
2. Postma DS, Rabe KF. The Asthma-COPD Overlap Syndrome. *N Engl J Med* 2015;373:1241-9.
3. Rennard SI, Locantore N, Delafont B, et al. Identification of five chronic obstructive pulmonary disease subgroups with different prognoses in the ECLIPSE cohort using cluster analysis. *Ann Am Thorac Soc* 2015;12:303-12.
4. Duyzend MH, Eichler EE. Genotype-first analysis of the 16p11.2 deletion defines a new type of "autism". *Biol Psychiatry* 2015;77:769-71.
5. Grouse L. Can epigenetic changes cause COPD? Canaries in the coal mines. *J Thorac Dis* 2013;5:80-1.
6. O'Roak BJ, Stessman HA, Boyle EA, Recurrent de novo mutations implicate novel genes underlying simplex autism risk. *Nat Commun* 2014;5:5595.
7. Hormozdiari F, Penn O, Borenstein E, et al. The discovery of integrated gene networks for autism and related disorders. *Genome Res* 2015;25:142-54.
8. Stessman HA, Bernier R, Eichler EE. A genotype-first approach to defining the subtypes of a complex disease. *Cell* 2014;156:872-7.
9. Claussnitzer M, Dankel SN, Kim KH, et al. FTO Obesity Variant Circuitry and Adipocyte Browning in Humans. *N Engl J Med* 2015;373:895-907.

Cite this article as: Grouse L. COPD diagnosis: genome first. *J Thorac Dis* 2015;7(10):E486-E488. doi: 10.3978/j.issn.2072-1439.2015.10.38