The Journal of Clinical Investigation

Clinical testing should be individualized, not based on populations

David R. Walt

J Clin Invest. 2019;129(9):3472-3473. https://doi.org/10.1172/JCI130903.

Viewpoint

Biomarker discovery is one of the fastest growing fields in clinical diagnostics. A biomarker is defined as "a biological molecule found in blood, other body fluids, or tissues that is a sign of a normal or abnormal process, or of a condition or disease" (1). Oncology, neurodegenerative disease, cardiovascular disease, and infectious disease fields are all exploring biomarkers for early detection, recurrence, and therapeutic monitoring. For example, discovering biomarkers for cancer using circulating tumor DNA is one of the most active areas, because tiny amounts of tumor DNA can be detected in blood by deep sequencing (2). Normal values are based on population measurements Traditional clinical laboratory tests have a normal range associated with each analyte measured. For example, nonfasting glucose should be in the range of 70–110 mg/dL, LDH in the range of 50–200 U/L, and cholesterol levels should be below 200 mg/dL. The ranges and recommended levels constituting "normal" have been determined by measuring thousands of subjects over time. Even so, we recognize that there is a problem with such absolutes. In response, we define borderline levels for which the values are near to the normal range but are not high or low enough to diagnose patients with disease. Prediabetes and borderline high cholesterol are good examples. In these cases, such borderline measurements may be useful in recommending […]

Find the latest version:

https://jci.me/130903/pdf

Clinical testing should be individualized, not based on populations

David R. Walt

Department of Pathology, Brigham and Women's Hospital, Boston, Massachusetts, USA. Harvard Medical School and Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, Massachusetts, USA.

Biomarker discovery is one of the fastest growing fields in clinical diagnostics. A biomarker is defined as "a biological molecule found in blood, other body fluids, or tissues that is a sign of a normal or abnormal process, or of a condition or disease" (1). Oncology, neurodegenerative disease, cardiovascular disease, and infectious disease fields are all exploring biomarkers for early detection, recurrence, and therapeutic monitoring. For example, discovering biomarkers for cancer using circulating tumor DNA is one of the most active areas, because tiny amounts of tumor DNA can be detected in blood by deep sequencing (2).

Normal values are based on population measurements

Traditional clinical laboratory tests have a normal range associated with each analyte measured. For example, nonfasting glucose should be in the range of 70–110 mg/dL, LDH in the range of 50–200 U/L, and cholesterol levels should be below 200 mg/dL. The ranges and recommended levels constituting "normal" have been determined by measuring thousands of subjects over time. Even so, we recognize that there is a problem with such absolutes. In response, we define borderline levels for which the values are near to the normal range but are not high or low enough to diagnose patients with disease. Prediabetes and borderline high cholesterol are good examples. In these cases, such borderline measurements may be useful in recommending lifestyle changes to potentially prevent later disease onset. In contrast, the meaning of absolute levels of other biomarkers is less clear cut. For example, prostate-specific antigen (PSA),

which is still used as a screening test for prostate cancer, levels above 4.0 ng/mL are considered above the normal range, and men who test above this level are recommended for biopsies. Abundant evidence indicates that PSA levels are highly individualized and affected largely by genetic predisposition (3). Predisposition for PSA blood levels is driven by genetic variation that likely affects promoters, inducers, repressors, and copy number variation that control gene expression and consequent protein expression (4). Some men are genetically predisposed to have higher levels of PSA, and those levels are stable over time. No change should mean no concern. Yet many physicians only look at the most recent results, and if the value is above the "normal" range then a biopsy is recommended. Clearly, clinical symptoms and physical exam findings are also used for diagnosis in addition to lab results, but it is important to individualize lab values beyond the "normal ranges" based on populations (Figure 1).

Baseline measurements are essential with more sensitive measurement technologies

With the advent of newer technologies that can measure exceedingly low concentrations of nucleic acids (mRNAs, microRNAs) and proteins, we are discovering new biomarker candidates. In many cases, these molecules vary over a hundred-fold in concentration and, in some cases, over ten thousand–fold between ostensibly healthy individuals (5). How is this variation biologically possible? Biology is robust, and when one component in a system is perturbed, compensatory

changes to maintain homeostasis occur in other pathways (6). For clinicians used to having normal ranges that vary by several-fold, it is exceedingly difficult to define "normal" when patient protein values can vary by over a thousand-fold. Additionally, it is possible that biomarker distribution in a population may be skewed rather than normally distributed. This variability means that sick individuals likely will have biomarker values that completely overlap with those of healthy controls. Disease will be noticeable only by observing changes in these values within individual patients.

Baseline measurements will become even more challenging as we begin to incorporate protein and nucleic acid panels into medical diagnosis. By measuring many different species simultaneously and coupling these panels with computational or machine learning algorithms (AI), it is possible to identify patterns of biomarkers where the absolute concentrations are unimportant but the relationships *between* biomarkers are diagnostic. Such panels, coupled with these algorithms, overturn conventional approaches, because, in most cases, the values of all the biomarkers constituting the panel fall within the range of healthy individuals. It is the *pattern* of biomarkers that distinguishes healthy from diseased individuals, not the particular values of a given marker.

Longitudinal sampling is required for precision diagnostics

We can no longer rely on normal ranges for many of these new biomarkers. Each person has a healthy baseline. Their present self should be compared to their previous healthy self rather than to a population average. Variations in baseline levels will need to be understood over time — what is the normal variability of a particular protein within an individual over the course of a day, week, year? How do baselines

Biomarker with narrow distribution range   Biomarker with wide distribution range
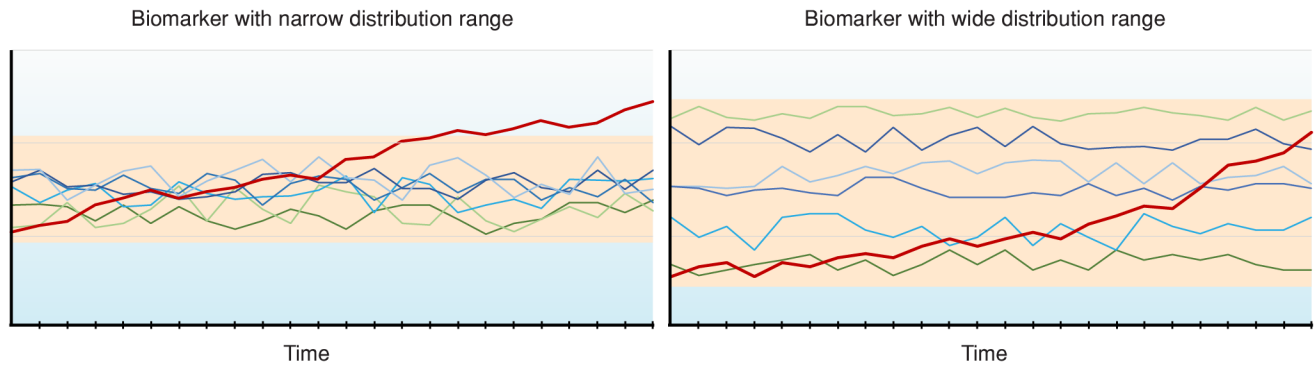
Time   Time

Figure 1. Hypothetical data set depicting biomarkers with a narrow or wide distribution range across patient populations. Left: The baseline levels of a biomarker vary over time in multiple individuals (each line represents a unique individual). The range of healthy baseline levels is narrow, as indicated by the shaded area (peach). One individual's biomarker concentration increases over time (red) and eventually goes outside the normal range, making it easy to detect an abnormal value. Right: The range of healthy baseline levels varies widely between patients. One individual's biomarker concentration increases significantly over time (red) but stays within the normal distribution range. For this individual, it is important to compare the biomarker concentration value to previous measurements to ascertain that the level exceeds this individual's baseline. Illustrated by Rachel Davidowitz.

change with age? We will need to understand these changes through studies of large populations and by using available well-annotated sample banks containing longitudinal samples of initially healthy individuals who are diagnosed with disease over time. Once we understand these variations, we will be able to apply true precision medicine where we diagnose patients based on changes from their former healthy selves.

Depending on the disease, different sampling frequencies will be required. In some cases, such as cancers, cardiovascular disease, and neurodegenerative diseases, there may be "sentinel" biomarkers that can be tested annually. These biomarkers may only be able to ascertain that something is changing, and additional testing with more specific biomarker panels may be necessary to provide a full diagnosis. In other cases, such as in individuals where there is genetic predisposition for disease, more frequent and comprehensive biomarker panels should be used. For therapeutic efficacy, one will need to measure biomarkers that demonstrate that the patient is returning to a more normal baseline. For example, reduced inflammation would be signaled by beneficial cytokine changes upon treatment with antiinflammatory agents. For recurrence monitoring, biomarkers should be measured that indi-

cate that the patient is moving away from a healthy state and returning to a disease state. Again, these changes would be individualized — what constitutes a healthier state for one individual might overlap with an unhealthy state for another.

In addition to the benefits to diagnosis, the information obtained by measuring biomarker values over time will likely lead to better mechanistic understanding of the pathobiology and disease pathways. In turn, improved understanding will lead to rational therapeutic interventions. By looking at high-resolution changes *within* an individual as they progress from good health to illness and then ideally return to their healthy baseline over time, rather than looking at differences from a population average, the perturbations should be clear and informative biomarkers can be identified.

The ideas proposed here will not lead to discoveries overnight because the underlying research will require a commitment by clinicians, researchers, and a large number of patients to procure blood, urine, and other samples over a long time. This approach differs dramatically from studies in which patients with particular diseases consent to provide a single blood sample and comparisons are made between healthy and diseased patients to identify the biomarkers that correlate with disease — the population approach. We need to

commit the time and resources necessary to identify the meaningful biomarkers that correlate with disease and health within individuals to truly realize the full potential of precision medicine.

Address correspondence to: David R. Walt, BWH-Pathology, 60 Fenwood Road, HBTM 8002J, Boston MA 02115. Phone: 857.307.1112; Email: dwalt@bwh. harvard.edu.

1. National Cancer Institute at the NIH. National Cancer Institute dictionary of cancer terms Website. http://www.cancer.gov/publications/dictionaries/cancer-terms/def/biomarker. Accessed August 14, 2019.
2. Aravanis AM, Lee M, Klausner RD. Next-generation sequencing of circulating tumor DNA for early cancer detection. *Cell*. 2017; 168(4): 571–574.
3. MacGregor AJ, Gallimore JR, Spector TD, Pepys MB. Genetic effects on baseline values of C-reactive protein and serum amyloid a protein: a comparison of monozygotic and dizygotic twins. *Clin Chem*. 2004;50(1):130–134.
4. Parikh H, et al. Fine mapping the KLK3 locus on chromosome 19q13.33 associated with prostate cancer susceptibility and PSA levels. *Hum Genet*. 2011;129(6): 675–685.
5. Wu D, Dinh TL, Bausk BP, Walt DR. Long-term measurements of human inflammatory cytokines reveal complex baseline variations between individuals. *Am J Pathol*. 2017;187(12):2620–2626.
6. Kitano H. Towards a theory of biological robustness. *Mol Syst Biol*. 2007;3:137.