

Biology, Social Environment, and Personalized Medicine

Ralph I. Horwitz^a Allison Hayes-Conroy^b Burton H. Singer^c

^aTemple Transformative Medicine Institute and ^bGeography and Urban Studies, Temple University, Philadelphia, PA, and ^cEmerging Pathogens Institute, University of Florida, Gainesville, FL, USA

Variability is the law of life and as no two faces are the same, so no two bodies are alike and no two individuals react alike and behave alike under the abnormal conditions which we know as disease.
Sir William Osler [1]

The individualized care of patients has always been the singular focus of physicians in the practice of medicine. The commitment of physicians to the care of sick patients has remained steadfast despite the urgings of health policy experts who often prefer an emphasis on population medicine or of health system leaders who emphasize system-level efficiency and economic performance.

For most of the history of medicine, clinical decision-making for individual patients was guided by the prevailing theory of diseases, the personal experience of the doctor treating similar patients in the past and the authoritative, if often misguided, recommendations of eminent clinicians. The Greco-Roman concept that health reflects a balance in the 4 humors endured for 2 millennia. It was not until the Renaissance that the humoral concept of disease was discarded and the fields of anatomy, chemistry and human physiology influenced the development of more science-based medical practice.

These advances in the physical and biological sciences provided much needed explanations for how the body functioned in health and disease. Leading scientists such as Claude Bernard developed the concept of internal physiological balance, later named homeostasis by Walter Cannon [2, 3]. Edward Jenner introduced the method of vaccination, and Joseph Lister demonstrated the value of antiseptics [4, 5]. It was not long before those discoveries and others (e.g., Landsteiner's system of blood compatibility, Banting and insulin for diabetes) fundamentally strengthened the benefit of treatment for individual patients.

Advances in the social sciences were notable too. Progress in social theory by Comte and Spencer argued for the importance of understanding social phenomena in health and disease. Advances were also made in psychology by Dewey (educational theory), by Watson (behavioral conditioning) and by those who contributed to the rise of the cognitive sciences. Ironically none of these advances in social or behavioral science were integrated into clinical medicine or materially affected the care of the individual patient by physicians educated largely in the biological sciences.

Leading scientists in the 19th century came to believe in the biological causes of disease and the “one agent, one dis-

ease” theory of etiology. An exception to this line of reasoning was exemplified by Rudolf Virchow who was a pre-eminent 19th century scientist and who contributed most notably to the understanding of disease at the cellular level. But Virchow also recognized that biology and social conditions were intimately linked together in the development of diseases. In 1848 while working as a staff physician at Royal Charité Hospital in Berlin, Virchow participated as a member of a commission to investigate an outbreak of typhus in the Prussian province of Upper Silesia. He meticulously described the disease and its distribution but also the social conditions of poverty and poor hygiene that contributed to the epidemic. Virchow emerged from this experience convinced that health and disease in populations reflect where people live, what they eat, the work they do, the air and water they consume, their connectedness with others and the status they occupy in the social order. But like others before and after him, Virchow viewed the biology for individuals and the social determinants for populations as separate from each other [6].

Probability versus Determinism

The separation between the science of individual biology and population level social and environmental conditions was mirrored in an emerging debate on the use of statistics in medicine. In 1835, when a French surgeon (Jean Civiale) introduced a new method of bladder stone removal, he used statistics to claim his method was superior to the older technologies. His claim touched off a furious debate leading Claude Bernard to make the following statement: “A great surgeon performs operations for stone, later he makes a statistical summary and concludes from these statistics that the mortality law for this operation is two out of five. Well I say that his ratio means literally nothing scientifically. What should be done instead of getting facts empirically is to study them more accurately, each in its special determinism. We must study cases of death with great care and try to discover in them the cause of mortal accidents so as to master the cause and avoid the accident” [7].

The controversy over Civiale’s claims and Bernard’s repudiation of his method prompted the French Academy of Science to establish a commission to adjudicate the dispute. The commission was led by the esteemed statistician Simeon Poisson who wrote as follows: “In statistical affairs the first care is to lose sight of the man in order to consider him as only a fraction of a species. It is necessary to strip him of his individuality, to eliminate all accidental

effects that individuality would have been able to introduce into the question. It is altogether different in the domain of medicine” [7].

A century later, when R.A. Fisher applied the method of randomization to compare the effects of different interventions in agricultural settings, he put into motion a process that led to the first randomized controlled trial (RCT) in medicine conducted by Austin Bradford Hill (the Medical Research Council study of the new antibiotic streptomycin in the treatment of tuberculosis) [8]. It was not long afterwards that the RCT became the accepted “gold standard” to determine whether a new or existing treatment worked. But despite the widespread embrace of the RCT by the medical community, those who knew the method best had the greatest reservations.

Group versus Individual

In his Heberden Oration in 1965 Austin Bradford Hill wrote: “This leads directly to a related criticism of the present controlled trial that it does not tell the doctor what he wants to know. It may be so constituted as to show without any doubt that treatment A is on the average better than treatment B. On the other hand that result does not answer the practicing doctor’s question, what is the most likely outcome when this drug is given to a particular patient?” [9].

Hill went on: “In a more sophisticated trial, can we identify the individual patient for whom one or the other of the treatments is the right answer? Clearly that is what we want to do and present day investigators ought to give far more attention to the problem. There are very few signs that they are doing so.”

Hill’s concerns were echoed 20 years later in a frequently cited paper by John Tukey who wrote on the problem of multiplicity in randomized trials. Speaking specifically about the application of trial results to clinical practice, Tukey wrote: “It is a difficult task to drive the nearly incompatible two horse team: On the one hand, knowledge of a most carefully evaluated kind where in particular questions of multiplicity are faced up to; and on the other, informed professional opinion, where impressions gained from statistically inadequate numbers of cases often, and so far as we can see, often should control the treatment of individual patients. The same physician or surgeon must be concerned with both what is his knowledge and what is his informed professional opinion as part of treating a single patient. I wish I understood better how to help in this essentially ambivalent task” [10].

Clinician's Dilemma and Role of Subgroups

The concerns of Bradford Hill and Tukey have framed the clinician's dilemma. The ideal information base for a clinician choosing treatment for a particular patient is an archive of clinical histories closely matched to the patient in question, where some individuals have received the contemplated treatment and others have not. Most of the data currently available about the results of using alternative treatments derive from RCTs. Unfortunately; the populations enrolled in such trials are far more heterogeneous than are required for a set of close matches to the patient being seen by the physician. Although these average treatment effects from heterogeneous RCTs may be acceptable for regulators licensing drugs, they are rarely sufficient for the clinician focusing on a single patient. This leads naturally to the clinician's dilemma: the clinician eagerly inquiring of the data about ever more refined patient subgroups and the RCT trialist strongly resisting such requests.

The trialist's resistance to analyzing subgroups was famously illustrated in a whimsical analysis of the ISIS-2 trial in which the overall benefit of treatment with aspirin therapy after acute myocardial infarction was absent in patient subgroups born under 2 of the 12 astrological signs [11]. The authors used this subgroup result to argue that subgroups are more likely to be misleading deviations from the overall average than indicators of meaningful interaction. The physician who relies on subgroups, it was argued, may harm their individual patient, since "undue emphasis in a particular subgroup may result in inappropriate treatment."

Certainly this is true if subgroups have no greater foundation for a differential clinical effect than astrological star signs. Although output from many clinical trials contains subgroup analyses, the results are presented with caveats about their interpretation, and the subgroups are formed using single features (e.g. age, sex, and race/ethnicity) and without more sophisticated profiles created with a deeper appreciation of clinical biology or the effect of social, behavioral, and environmental influences.

Biology, Social Science, and Subgroup Analyses

Just when it appeared that trialists and clinicians had reached an irreconcilable impasse, advances in the biological and social sciences turned the debate in a new direction. Diffuse large B-cell lymphoma (DLBCL) is the most common subtype of non-Hodgkin lymphoma. The

disease is clinically heterogeneous. R-CHOP¹ induces complete remission in 75% of patients but long-term event-free survival is only 50–60%, and 30–40% of the patients eventually die from their disease. It is not surprising then that there has been intense interest in identifying patients with different long-term prognoses.

In 1993, clinical experts developed an international prognostic index that included age (>60 years), advanced stage disease (III, IV), and elevated lactate dehydrogenase, a performance status of 2 or higher and the presence of more than 1 site of extranodal disease. When patients were partitioned according to the number of adverse prognostic features, overall 10-year survival showed a gradient from low to high [12].

In 2000, Alizadeh et al. [13] proposed that the variability in clinical outcomes of DLBCL reflected previously unrecognized molecular heterogeneity in the tumors. Using DNA microarrays, they showed diversity in gene expression among the tumors of DLBCL patients that reflected variation in tumor proliferation rate, host response and the differentiation state of the tumors. Two molecularly distinct forms of DLBCL were identified. Patients with germinal B-like DLBCL had a significantly better overall survival than those with activated B-like DLBCL. Furthermore, when those molecular features were combined with clinical stages in the IPL, further clinically important distinctions in prognosis were demonstrated [13].

For many authorities, the molecular analysis of DLBCL ushered in the era of personalized medicine. Personalized medicine is often described as genomics-based knowledge that "promises the ability to approach each patient as the biological individual he or she is." The molecular distinctions in DLBCL and their impact on prognosis appear to validate this definition. However, the analyses by Alizadeh et al. also suggest that patient profiles that determine prognosis require both biological and clinical features. What about social and environmental characteristics?

The importance of these nongenomic factors was brilliantly described in a 2010 paper that studied the genetically homogeneous Berbers of Morocco to illustrate the interaction of biology, culture, and environment. The Souss region in southern Morocco contains several million people of 2 dominant ethnicities living in either

¹ R-CHOP is a chemotherapy where the initials correspond to the drugs used in the treatment. R, rituximab; C, cyclophosphamide; H, doxorubicin (hydroxydaunomycin); O, vincristine (Oncovin); P, prednisolone (a steroid).

cities or rural villages. The Amazigh Berbers are descended from the first modern humans who lived in North Africa 35,000 years ago and many of whom continue to live in traditional villages. The Arabs moved into southern Morocco between the 7th and 11th centuries and occupy lowland villages while both groups inhabit the cities.

The investigators collected peripheral blood samples from 284 healthy adults in 2008 from 4 locations including nearly equal numbers of men and women and of Amazigh and Arabs. Half of the samples were from 2 high-density urban communities and the other half from 2 rural villages. When gene expressions in peripheral white blood cells were examined from nomadic, mountain agrarian and coastal urban individuals, up to one third of the “transcriptome” was associated with different environments (diet, exposure to microbial organisms, and environmental stresses among other influences) in these genetically homogeneous people [14].

Biology, Behaviors, and Psychosocial Influences

Just as remarkably, psychosocial features of patients at the individual level also directly affect clinical outcomes. The relationship of treatment adherence to mortality after a myocardial infarction was studied among 2,175 participants in the Beta Blocker Heart Attack Trial that had data for measures of treatment adherence, clinical severity, and the psychological and social features that may influence postinfarction mortality. Overall, patients who did not adhere well to treatment (i.e., took less than 75% of prescribed medication) were 2.6 times more likely than good adherers to die during the follow-up. Poor adherers had an increased risk of death whether they were on β -blocker (OR = 3.1) or placebo (OR = 2.5). The increased risk of death for poor adherers persisted after adjustment for myocardial infarction severity, sociodemographic features, smoking, and such psychological characteristics as high life stress and social isolation [15].

All these observations begin to point to a new way of conducting clinical research so that it conforms more closely to clinical practice. Rather than reliance solely on the average results of groups of patients drawn exclusively from randomized trials, we can use information from biology, behavior, and psychosocial influences to describe distinct profiles of individual patients. Patient profiles would then be formed by integrating information across all of these domains, biology, behavior, and psychosocial.

To accomplish this objective requires improved taxonomic systems and improved measurement tools. In this regard, our colleagues in the biological sciences have advanced further and faster. For example a recent report in *Nature* described the narcissism of one person’s “integrative personal -omics profile” by merging the genomic sequence with RNA, protein, metabolic, and autoantibody profiles measured 20 times over a 14-month period. Not only were new disease susceptibilities identified, but the patient’s blood glucose levels escalated following a viral infection indicating a diabetic state that was only diagnosed later. A criticism of the paper was that it is anecdotal about 1 person. But that is also its strength according to geneticist George Church of Harvard Medical School. He went on to say that “medicine treats people as an individual and we should be conducting research with a similar attitude” [16].

Clinical medicine is fast developing new ways to profile patients. Current devices are capable of “digitizing” a person with wearable sensors that quantify physiological metrics such as vital signs, provide high-definition images of a person’s anatomy, and characterize the microbiome. Food and Drug Administration-approved smartphone electrocardiogram sensors are already available to patients along with immediate computer-assisted readers. It is expected that patients with hypertension and diabetes will soon be able to continuously monitor blood pressure and blood glucose via a smart watch, and that much of the physical exam will be performed remotely with digital devices.

Fava et al. [17] have provided a new roadmap to integrate biobehavioral and socioeconomic data into the emerging approach to individualized medical care. They point out that psychosomatic research has advanced to enable studies of complex biopsychosocial phenomena that may provide new effective modalities of patient care. They emphasize especially psychosocial variables affecting individual vulnerability, clinical course, and outcome of illness and the role that allostatic load (chronic stress) may play.

What remains missing, however, is a comprehensive approach that seeks to understand how the interactions among genetics, mind, body, behavior, and the environment affect both the risk for disease and the response to treatment. To date there are collections of data with catalogs of genetics and health records, but no data set has yet examined the holistic pattern of human behavior and environment in a defined group of subjects.

Next Steps towards Individualized Patient Care

The new era of individualized patient care will require new ways of thinking, new types of data that describe clinical, social, and behavioral features of individuals, and new methods of analysis. What will no longer suffice is the sole reliance on the average results from randomized controlled trials to guide treatment decisions. As emphasized earlier, these results do not answer the physician's question, namely: what is the likely effect of a given treatment on an individual patient at hand?

The first step in new ways of thinking is to acknowledge that the primary focus of the research and evidence generation must be at the individual level, rather than the group or population. The starting point for any clinically applicable analysis is a detailed characterization of an individual patient. This seemingly innocuous recommendation is likely to meet great resistance. Clinical research has little experience and less acceptance of studies oriented to individual level design and analysis. Claims of "bias" almost always ensue whenever such studies are suggested and descriptions of patients' clinical course and outcome are regularly disregarded if the data do not emerge from RCTs. So what exactly do we mean by individual level data, and how will it be utilized?

Individual level data will emphasize detailed and repeated measures over time of each subject across multiple domains including deep biology, clinical physiology, social and behavioral features, and the associated environment. It will then be necessary to create individual-based "profiles" that integrate the data across these domains and to assemble large archives of these profiles to enable the identification of approximate matches for each subject. In a word, the clinician seeking guidance for decision-making on management of an individual patient starts with a profile of that patient. She then asks how approximate matches to that patient have responded to a contemplated treatment and compares this with the histories of approximate matches who did not receive the treatment.

Some may immediately object to the complexity of the clinical experience of each patient and the density of data that will be required. Indeed, it is a legitimate worry that very large amounts of ("big") data may make it hard to find a signal amidst all of the potential noise. Our colleagues in the molecular and genetic sciences have successfully embraced the complexity of cellular and molecular pathways as they occur in an individual patient, and so must we in the clinical sciences who have somewhat more complex pathways to characterize.

Next, it is essential that we expand the customary data beyond the usual information collected in RCTs to include genomic and other "-omics" data, the microbiome, and the newly available data made possible through new digital technologies. These capabilities already exist. What does not yet exist are analytical methods that create patient profiles with a full integration of multiple domains and with repeated measures over time. Some initial promising steps are emerging however.

Tatonetti et al. [18] describe a novel strategy to identify adverse drug reactions using an adaptive data-driven approach that compensates for the missing clinical information that often hampers the use of large collections of data. The authors validated 47 of the known drug class interactions in an independent analysis of patient electronic health records and show that intricate analyses of observational clinical data can improve physicians' ability to predict as yet uncharacterized adverse drug effects and interactions. Although adverse drug reactions are not our primary focus, the methodology of the paper by Tatonetti et al. is relevant for the profile constructions that interest us.

A second example of innovative new approaches is Observational Health Data Sciences and Informatics (OHDSI, pronounced Odyssey) [19]. The OHDSI network envisions a future where observational studies will provide practice-based evidence using the previously unprecedented amount of available patient data and the use of computerized systems to process the data. What still remains as challenges, however, are the collection prospectively of highly personalized information on each patient and methods for creating individual and longitudinal patient profiles.

The data for these individualized profiles will need initially to be assembled from whatever archives enable their formation, including RCTs, cohorts, patient registries, electronic health records, administrative claims, but also financial records, consumer data and geographic locations. Later, data that enable individualized care will be collected prospectively and stored in archives that will be accessible to practicing clinicians at their personal computers.

The strategy of analysis will not proceed from large populations to smaller subgroups as occurs now. It starts with the individual patient, and uses his/her characteristics to define populations relevant to guide patient management – i.e. the approximate matches to the individual patient. Personalized medicine, so constituted, will require a new kind of evidence base containing the multidimensional patient histories associated with a large

number of common chronic disorders. The clinician treating an individual patient will be able to match her patient to all similar patients in the archive, including the treatment decisions and outcomes, as proposed by Feinstein et al. [20], who imagined a “library of clinical experience” to obtain personalized prognosis. Indeed, in 1972 Feinstein et al. put forth an early illustration of the process we have outlined and did so with far more limited computer capabilities than are currently available. The basic idea of starting with a profile of an individual patient and identifying approximate matches to the patient in a library of clinical experience and using these cases as a guide to decision-making for the patient at hand is in this visionary paper. In the short run, there is a pressing need to develop algorithms for approximate matching and to

do so on patients with a diversity of diseases. The matching would be based on much more comprehensive information than in the original example of Feinstein et al., but their fundamental philosophy about the character of the evidence needed to guide decision-making for individual patients is salient today.

With this background at hand, it is finally important to emphasize that advances that are occurring in molecular biology and human genetics must now occur in the clinical sciences as well. We are on the cusp of a great and exciting change in personalized clinical care. Fava et al. [17] remind us that these advances can only occur with a new understanding of the unity of mind (biography) and body (biology) that are reflected in the goals of psychosomatic medicine.

References

- Osler W: On the educational value of the medical society. *Yale Med J* 1993;IX:325.
- Gross CG: Claude Bernard and the constancy of the internal environment. *Neuroscientist* 1998;4:380–385.
- Cannon WB: Organization for physiological homeostasis. *Physiol Rev* 1929;9:399–431.
- Lakhani S: Early clinical pathologists: Edward Jenner (1749–1823). *J Clin Pathol* 1992;45:756–758.
- Lister BJ: On the antiseptic principle in the practice of surgery. *Clin Orthop Relat Res* 2010;468:2012–2016.
- Eisenberg L: Does social medicine still matter in an era of molecular medicine? *J Urban Health* 1999;76:164–175.
- Matthews JR: Commentary: the Paris Academy of Science report on Jean Cuvier's statistical research and the 19th century background to evidence-based medicine. *Int J Epidemiol* 2001;30:1249–1250.
- Medical Research Council: Streptomycin treatment of tuberculous meningitis. *Lancet* 1948;i:582–596.
- Hill AB: Reflections on controlled trial. *Ann Rheum Dis* 1996;25:107.
- Tukey JW: Some thoughts on clinical trials, especially problems of multiplicity. *Science* 1977;198:679–684.
- ISIS-2 (Second International Study of Infarct Survival) Collaboration Group: Randomized trial of intravenous streptokinase, oral aspirin, both, or neither among 17,187 cases of suspected acute myocardial infarction: ISIS-2. *Lancet* 1988;332:349–360.
- The International Non-Hodgkin's Lymphoma Prognostic Factors Project: A predictive model for aggressive non-Hodgkin's lymphoma. *N Engl J Med* 1993;329:987–994.
- Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, Boldrick JC, Sabet H, Tran T, Yu X, Powell JI, Yang L, Marti GE, Moore T, Hudson J Jr, Lu L, Lewis DB, Tibshirani R, Sherlock G, Chan WC, Greiner TC, Weisenburger DD, Armitage JO, Warnke R, Levy R, Wilson W, Grever MR, Byrd JC, Botstein D, Brown PO, Staudt LM: Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 2000;403:503–511.
- Idaghdour Y, Czika W, Shianna KV, Lee SH, Visscher PM, Martin HC, Miclaus K, Jadallah SJ, Goldstein DB, Wolfinger RD, Gibson G: Geographical genomics of human leukocyte gene expression variation in southern Morocco. *Nat Genet* 2010;42:62–67.
- Horwitz RI, Viscoli CM, Berkman L, Donaldson RM, Horwitz SM, Murray CJ, Ransohoff DF, Sindelar J: Treatment adherence and risk of death after a myocardial infarction. *Lancet* 1990;336:542–545.
- Dennis C: The rise of the narcissome. *Nature* 2012, DOI: 10.1038/nature.2012.10240.
- Fava GA, Cosci F, Sonino N: Current psychosomatic practice. *Psychother Psychosom* 2017;86:13–30.
- Tatonetti NP, Ye PP, Daneshjou R, Altman RB: Data-driven prediction of drug effects and interactions. *Sci Transl Med* 2012;125:125–131.
- Hripcsak G, Duke JD, Shah NH, Reich CG, Huser V, Schuemie MJ, Suchard MA, Park RW, Wong IC, Rijnbeek PR, van der Lei J, Pratt N, Norén GN, Li YC, Stang PE, Madigan D, Ryan PB: Observational Health Data Sciences and Informatics (OHDSI): opportunities for observational researchers. *Stud Health Technol Inform* 2015;216:574–578.
- Feinstein AR, Rubinstein JF, Ramshaw WA: Estimating prognosis with the aid of a conversational-mode computer program. *Ann Intern Med* 1972;76:911–921.