GENOMICS

What’s a Genome Worth?

Isaac S. Kohane* and Jay Shendure**

A recent study (Roberts et al.) explores considerations in estimating the current and potential clinical utility of whole-genome sequencing for individual patients.

Both in the public eye as well as within the scientific community (1), much of the current enthusiasm for genomics rests on the prospect that assaying a patient’s genetic information will improve the effectiveness of health care—that is, it will directly inform their medical care in ways that improve outcomes and reduce costs. However, it remains controversial whether these high hopes for what is termed “genomic medicine” are well founded. In this issue of Science Translational Medicine, Roberts et al. (2) describe their use of epidemiological data from twin registries to attempt to answer a pressing question in genomics: How much predictive value for disease risk will actually be obtained when the genomes of healthy individuals are routinely sequenced in a clinical setting?

This is a timely topic not only because a week does not pass without the publication of a diagnostic result obtained through whole-genome sequencing but also because the price of whole-genome sequencing has decreased sufficiently that it is now comparable to the costs of other high-tech diagnostic tests such as high-resolution cranial magnetic resonance imaging. The convergence of increasing utility and plummeting costs highlights the relevance of the specific question posed by Roberts et al. (2): What is the maximum potential benefit of whole-genome sequencing with regard to the ability to predict future diseases? We can refer to this predictive task as the quantification of the lifelong risk for a specific disease in an asymptomatic individual without any clinical evidence of disease.

To address this question, Roberts et al. leverage the overall disease prevalence and the rates of concordance between monozygotic twins for 24 common diseases, including coronary heart disease, Alzheimer’s disease, and breast cancer. An analysis by Roberts et al. (2) suggests that taking this approach might typically reveal a significantly increased risk for the individual for only one out of 24 common diseases.

Fig. 1. Insights from genomics: From the cradle to the grave. With the cost of whole-genome sequencing plummeting, the DNA in a blood sample obtained at a single time point (e.g., at birth) could be sequenced and deposited in a database accessible to the individual and designated clinicians. This whole-genome sequencing data could be used prognostically, while the individual is asymptomatic, to calculate the lifetime risk of that person developing common diseases, such as coronary heart disease, Alzheimer’s disease, and breast cancer. An analysis by Roberts et al. (2) suggests that taking this approach might typically reveal a significantly increased risk for the individual for only one out of 24 common diseases. However, it may turn out that the real value of whole-genome sequencing in the clinic lies in areas other than asymptomatic prognosis—for example, the precise diagnosis of a disease at presentation or the selection of an appropriate therapy.

United States. They introduce the concept of “genometypes”—groups of genomes conferring an identical genetic risk. Because monozygotic twins essentially share a genome, they, by definition, reside within the same genometype. However, only a subset of all possible distributions of genetic risk of disease is compatible with the epidemiological data. For example, the prevalence and monozygotic twin concordance rate of a given disease might equally well be explained by a small fraction of genometypes conferring a high genetic risk or a larger fraction of genometypes conferring a modest genetic risk. Roberts et al. assume a best-case scenario—that is, the distribution of genometypes that is both compatible with the epidemiological data and maximizes the clinical utility of whole-genome sequencing. Assuming that there will come a day when genomics researchers will be able to comprehensively decode the heritability signal present in whole genomes, the result is an upper bound of the performance of whole-genome sequencing for the prognosis of common diseases among asymptomatic individuals.

Given the aspirations of the many teams engaged in this area, most notably the direct-to-consumer genomics companies of the past decade, the results may appear disappointing at first glance. For the majority of the diseases tested (23 out of 24), most individuals would receive negative test results and these negative results would not meaningfully decrease their estimated risk for developing that disease. However, this perspective obscures a more positive result: 90% of tested individuals would in fact have at least one disease for which they were predicted to have a higher risk at a threshold that was selected to be of clinical utility. In other words, although a whole-genome sequence is highly unlikely to serve as a crystal ball across most diseases for most patients, it still has value by identifying subsets of patients that are at a clinically significant increased risk for specific diseases.

Unfortunately, we remain far away from being able to reliably interpret the entirety of the estimated heritability for common diseases encoded by variation in the human genome (3). Nearly all of the millions of genetic polymorphisms observed to date remain of unknown clinical significance. Moreover, many of the variants for which there is published evidence for impact on disease risk may not confer the same risk in an asymptomatic population of interest as they do in the population where they were originally implicated. Consequently, during
the interval in which we sort out which vari-
ments and combinations thereof contribute to
disease risk, not only will whole-genome se-
quencing fall short of the upper bound esti-
mates for asymptomatic prognosis provided
ted by Roberts et al., but also there may be sub-
stantial violation of the medical imperative
to do no harm. Specifically, there is the threat of
the “incidentalome”—that is, the likely
proliferation of false positive incidental find-
ings consequent to performing genetic test-
ning on a genome-wide scale (4).

Notwithstanding the results of Roberts et al.,
there are an increasing number of pub-
lished examples of whole-genome sequenc-
ing demonstrating its value in specific clini-
cal scenarios. These include whole-genome
or exome sequencing that has facilitated
unanticipated diagnoses of Mendelian dis-
orders (5–7), as well as aiding in therapeutic
decision-making for cancer patients (8). For
example, Worthey et al. reported the lifesav-
ing diagnosis of a rare but treatable Mendel-
lian disorder, X-linked inhibitor of apoptosis
deficiency, in a pediatric patient by exome
sequencing (5). Meanwhile, Jones et al. have
used genome-wide mutational analysis to
guide the selection of specific kinase inhibi-
tors to treat a patient with adenocarcinoma
of the tongue, a rare cancer with no estab-
lished treatment protocol (8). Furthermore,
the large catalog of non–whole-genome se-
quencing genetic tests (i.e., targeted to spe-
cific variants, genes, or gene panels) that are
already a part of medicine is clear evidence
that genetic information can and does add
value to medical care. The broad set of clini-
cal tasks that might well benefit from whole-
genome sequencing that are not addressed by
Roberts et al. include genetic testing for all
aspects of reproductive health. This includes
preconception carrier screening to detect,
for example, the Tay Sachs disease mutation,
preimplantation genetic diagnosis to select
healthy embryos, prenatal diagnostic testing,
and newborn screening to definitively diag-
nose all Mendelian disorders at birth (Fig. 1).
In the context of cancer, whole-genome se-
quencing could be used to provide a rational
selection of targeted therapeutics based on
the precise molecular pathways disrupted in
an individual patient’s tumor and also to
monitor the evolution of the tumor in re-
sponse to those treatments.

The differences between these clinical sce-
narios and that considered by Roberts et al.
are manifold. But a key difference is the fact
that the probabilistic dependencies invoked
while interpreting and acting on genetic in-
formations in the context of a specific clinical
situation are very different from those when
the same task is carried out in an asympto-
matie population. These clinical scenarios, rath-
er than prognosis in an asymptomatic popu-
lation, may well end up becoming the raison
d’être of clinical whole-genome sequencing.
However, it is important to recognize that the
value of whole-genome sequencing in these
and other contexts cannot be estimated from
the model that Roberts et al. have developed.

Could the analysis of Roberts et al. have
significantly underestimated the value of
whole-genome sequencing even for the task
of asymptomatic prognosis? As the authors
themselves recognize, there are some limita-
tions of their study that cannot be fully
addressed at this time. First, several of the
24 diseases considered may not be inde-
pendent—for example, a diagnosis of breast
cancer increases one’s risk of developing
ovarian cancer. Such dependencies would
likely alter the informativeness of genetic
profiling. Also, the populations in which
many of the twin studies were conducted,
such as those in Sweden and Finland, are
considerably less heterogeneous than those
in other countries, such as the United States.
This might affect some of the prevalence and
heritability estimates that the current study
is based upon. Lastly, a single (and neces-
sarily somewhat arbitrary) definition of clinical
utility is applied by Roberts et al. on a per-
individual basis. However, it may be the case
that whole-genome sequencing is extremely
useful (i.e., lifesaving) for some genotypes,
such that the “average utility” makes its use
for asymptomatic prognosis worthwhile.

How much value must whole-genome se-
quencing add in order for it to make sense
to introduce it into routine clinical care? A
back-of-the-envelope estimate suggests that
the bar is remarkably low. Let us assume that
the all-inclusive cost for ordering a “genome
test” for a newborn in the United States drops
to $1,000. Given that one’s germline genome is
essentially static (and a genome sequenced
at birth can effectively be consulted through-
out one’s lifetime), this cost can be amortized
over the individual’s 78-year life expectancy
for an effective cost of just $13 per year.
In the United States, we currently spend about
$9,000 per capita per year on health care, to
which this $13 would add about a tenth of
one percent. As the cost of whole-genome
sequencing may well drop even lower than
$1,000 within the next few years, the “value
added” threshold that whole-genome se-
quencing must achieve to justify its wide-
spread implementation may be even lower.

It is a measure of the rapid maturation of
this field that even as we are still sorting out
the genetic basis of human disease, we are
already able to make informed, if tentative,
estimates of the potential clinical utility of
whole-genome sequencing as a routine test
that will be applied to every citizen. Many
more refinements of these estimates will
have to be elaborated if we are, as a society,
going to make an informed decision as to
when and for whom to incorporate whole-
genome sequencing into the routine practice
of clinical medicine.

REFERENCES

1. E. D. Green, M. S. Guyer; National Human Genome Re-
search Institute, Charting a course for genomic medicine

2. N. J. Roberts, J. T. Vogelstein, G. Parmigiani, K. W. Kinzler,
B. Vogelstein, V. E. Velculescu, The predictive capac-
ity of personal genome sequencing. Sci. Transl. Med. 4,
133ra58 (2012).

A. Hindorff, D. J. Hunter, M. I. McCarthy, E. M. Ramos, L.
R. Cardon, A. Chakravarti, J. H. Cho, A. E. Guttmacher,
A. Kong, L. Kruglyak, E. Mardis, C. N. Rotimi, M. Slatkin,
D. Valle, A. S. Whittemore, M. Boehnke, A. G. Clark, E. E.
Eichler, G. Gibson, J. L. Haines, T. F. Mackay, S. A. McCarr
roll, P. M. Visscher, Finding the missing heritability of

4. I. S. Kohane, D. R. Masys, R. B. Altman, The incidentalome:
A threat to genomic medicine. JAMA 296, 212–215

5. E. A. Worthey, A. N. Mayer, G. D. Syverson, D. Helbling, B.
B. Bonacci, B. Decker, J. M. Serpe, T. Dasu, M. R. Tschann,
R. L. Veith, M. J. Basehore, U. Broeckel, A. Tomita-
Mitchell, M. J. Arca, J. T. Casper, D. A. Margolis, D. P. Bick,
M. J. Hessner, J. M. Routes, J. W. Verbsky, H. J. Jacob, D.
P. Dimmock, Making a definitive diagnosis: Successful clini-
cal application of whole exome sequencing in a child
with intractable inflammatory bowel disease. Genet.

A. Nayir, A. Bakkaloglu, S. Ozen, S. Sanjad, C. Nelson-
Williams, A. Fehr, S. Mane, R. P. Liffman, Genetic diagnosis
by whole exome capture and massively parallel DNA se-
(2009).

7. M. N. Bainbridge, W. Wisniewski, D. R. McDoulc, J. Fried-
man, C. Gonzaga-Jaureguin, I. Newsham, J. G. Reid, J.
K. Fink, M. B. Morgan, C. M. Gingras, D. M. Muzny, L. D.
Hoang, S. Yousaf, J. R. Lupsic, R. A. Gibbs, Whole-genome
sequencing for optimized patient management. Sci.

8. S. J. Jones, J. Laskin, Y. Y. Li, O. L. Griffith, J. An, M. Bills
ken, Y. S. Butterfield, T. Ceard, E. Chua, R. Corbett, A. P.
Fejes, M. Griffith, J. Yee, M. Martin, M. Mayo, N. Melnyk, R.
D. Mor-
rin, T. J. Pugh, T. Severson, S. P. Shah, M. Sutcliffe, A. Tam, J.
Terry, N. Thieness, T. Thomson, R. Varhol, T. Zeng, Y. Zhao,
R. A. Moore, D. G. Huntsman, I. Birol, M. Hirst, R. A. Holt,
M. A. Marra, Evolution of an adenocarcinoma in response to
selection by targeted kinase inhibitors. Genome Biol.
11, R82 (2010).

Citation: I. S. Kohane, J. Shendure, What’s a genome worth?
