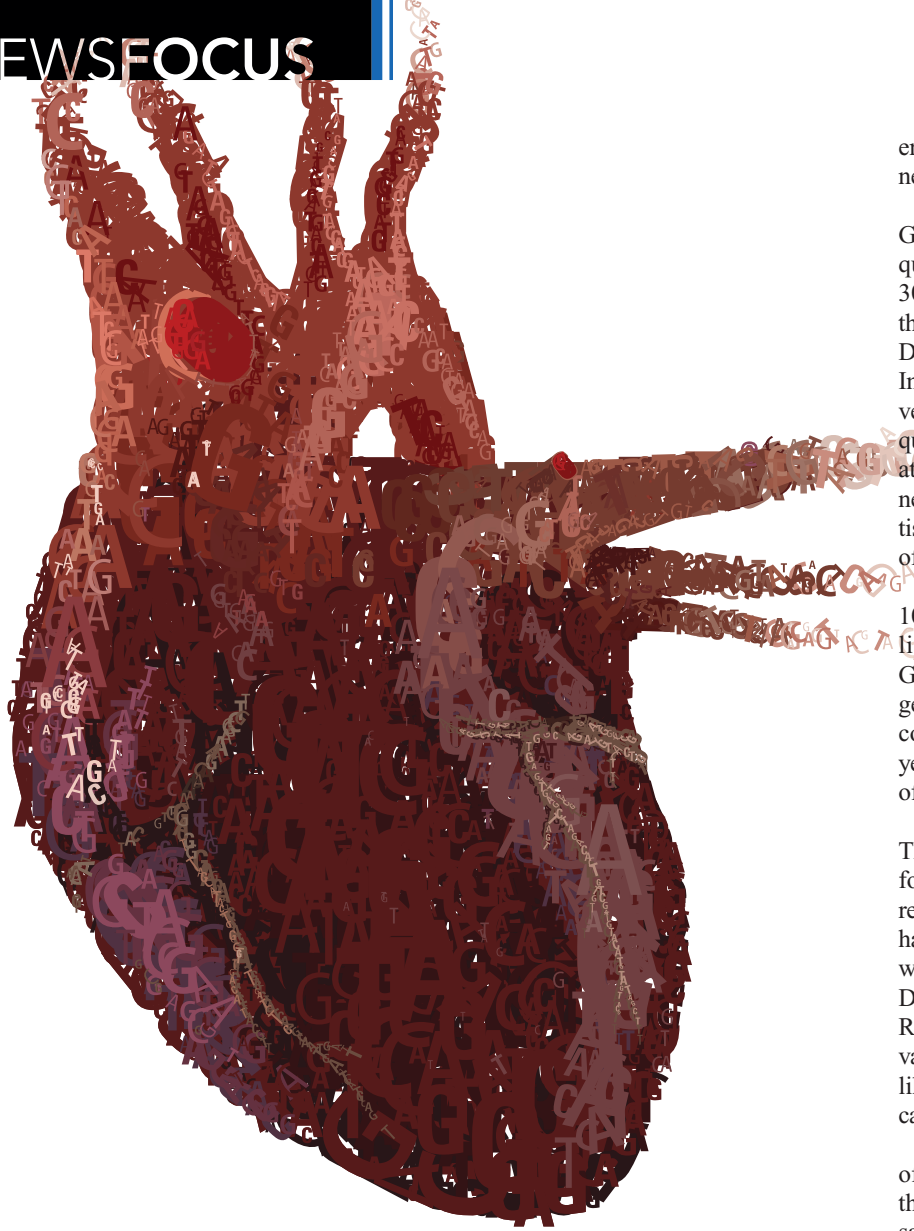
Table 2 Grading the credibility of the evidence for individual gene–disease associations: some proposed grading criteria and their limitations in interpreting recurring weak associations

Axis	Proposed grading[a]	Comments
Effect size	Small effect size (RR<2) has lowest grade while large effect size considered best (RR>5)	Most biologically causal factors are expected to have RR < 2. Many may be beyond the limit of analytical ability
Amount of evidence/replication	Single or few scattered studies have lowest grade while large-scale inclusive analyses are best	The more information the better the inference, although it may be difficult to set hard rules for the amount of replication for weak associations. There is a risk for endless replication
Protection from bias	Clear presence of bias gets poor grade while clear strong protection gets high grade	Most studies will be in between. Absolute protection from bias is hard to achieve. More empirical evidence and consensus is needed on which biases are more serious than others.
Biological plausibility	No functional data scores lowest while convincing biological data scores highest	Need consensus and empirical evidence for the importance of specific items of biological plausibility
Relevance	Graded according to clinical or public health application	Individual weak associations will have little relevance to use for genetic testing because of their poor predictive ability especially for rare conditions

[a] Grading proposed by Ioannidis.[57]



CREDIT: M. TWOMBLY/*SCIENCE*

Major Heart Disease Genes Prove Elusive

So far, genome-wide association studies have not found common genes with a big impact on heart health; researchers hope that the low-effect genes they are finding will help identify pathways and drug targets

THE EXCITEMENT BEGAN 5 YEARS AGO, when a study of 146 Caucasian volunteers turned up a common gene variant among those with the eye disease macular degeneration. Researchers had used a new strategy: They scanned large stretches of the genomes of the sick and the healthy and found a single DNA base that was much more likely to be present in those whose eyes were failing.

The finding was remarkable: Relatively few people participated in the study, yet those with two copies of the suspect gene variant had 10 times the risk of macular degeneration, a huge increase. Furthermore, the method the group used, called genome-wide association (GWA), had some big advantages: It was unbiased, testing thousands of gene-disease associations at once, not just a researcher's favorites. And it pointed to common variants, found in at least 5% of individuals studied. GWA studies offered hope of identifying people at risk for diseases, uncovering new disease mechanisms, and finding new targets for therapy.

Almost immediately, researchers applied GWA to other conditions. But they were quickly stymied. "People did studies with 300 or 500 people and didn't find anything, then did 1000 and didn't find anything," says Deepak Srivastava, who directs the Gladstone Institute of Cardiovascular Disease at the University of California (UC), San Francisco. It quickly became clear that macular degeneration was an exception. Most GWA studies needed 10,000 or more volunteers to get a statistically significant result, because the effect of each gene was so small.

Since the human genome was sequenced 10 years ago, technology has moved with lightning speed; many now believe that GWA methods, which cover a fraction of the genome, are becoming obsolete. Sequencing costs continue to plunge, and within a few years sequencing entire genomes of hundreds of subjects will be financially feasible.

What has the GWA experience taught us? The results from one group of GWA studies, for heart disease, are typical, with a mixed record and an uncertain legacy. The technique has identified dozens of variants, but all have weak effects; so far, almost none has led to DNA changes that actually cause disease. Researchers have had more success finding variants that link to tightly defined conditions like high cholesterol than to heart failure, a catch-all disease.

"At the end of the day, we have a bunch of loci and genes, but none of them" do all that much to raise the risk of heart disease, says Eric Topol, a cardiologist and director of the Scripps Translational Science Institute in San Diego, California. Nor have they yet altered our understanding of how the heart falters—knowledge, Topol says, that will take time to develop.

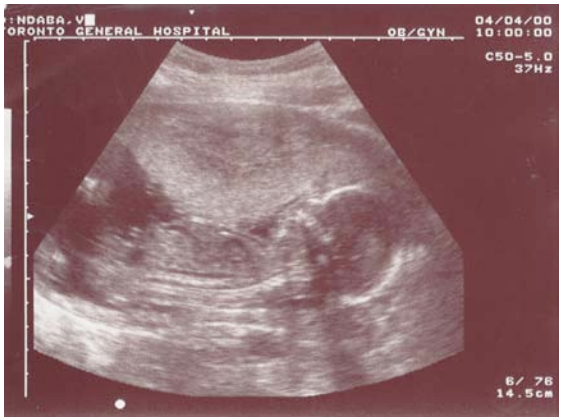
GWA studies still have many backers. "We have new technology that's enabled us to look at things we've never seen before," says Bruce Psaty, a cardiovascular disease epidemiologist at the University of Washington (UW) School of Medicine in Seattle. And Francis Collins, director of the National Institutes of Health (NIH), has said that the approach has provided "1000 new drug targets" (*Science*, 28 May, p. 1090).

Clues missing

The first GWA results for heart disease hit in 2007. Three studies examined coronary artery disease, in which plaque builds up in the arteries and narrows them. Together with subsequent studies, they identified 12 new genetic variants, called single-

Published by AAAS

Why study gene-environment interactions?

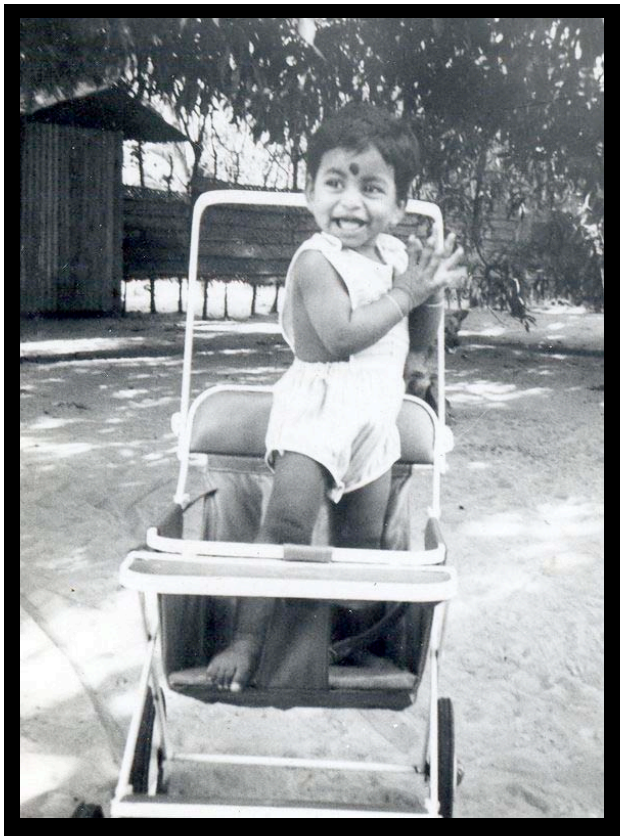


- Most disease burden is jointly determined by interaction of individual genetic endowments and complex sequence of environmental factors
- These gene-environment interactions require decades to fully manifest over the life course
- Diseases and conditions of later life occur in some and not others because of intense interactions between particular genetic constitutions and particular sequence of social and physical environments

CIHR IRSC

Canada

Why study gene-environment interactions? cont'd



- BUT…little is known about underlying causes of these conditions and why they are now increasing in frequency – for e.g. asthma
- Requires study of these sequential events in large numbers of people over time, on whom baseline genetic and repeated environmental exposures are taken, to:
  - understand the causal pathways; and,
  - develop disease prevention strategies

Studying Genetic and Environmental Contributions to Disease Causation: An Uneven Playing Field

Measurement Attribute	Genetic Exposure Measures	Environmental Exposure Measures
Time-varying?	No – one sample per lifetime is enough (unless gene expression arrays are used)	Yes – new samples needed whenever exposure changes
Data Collection Costs	Cheap (on a sample)	Expensive (real-time assays)
Sample Storage (for later analysis)	Easy (buccal swab, buffy coat)	Difficult (e.g. air/water/diet samples)
Data Analysis Costs	Getting cheaper by the day	Getting Costlier (as awareness of chemical/physical/biological complexity increases)
Overall Ease & Cost of Accurate Ascertainment	**Easy / Cheap**	**Difficult / Costly**

Comparison of "Huge, Data-Thin" Cohorts (e.g. U.K. BioBank) And "Small, Data-Thick" Cohorts (e.g. Southampton)

Cohort Attribute	Huge – Thin	Small – Thick
Cost Per Subject due to:	Low (e.g. < $500. / data-wave)	High (if > $1,000. / data-wave)
Sample Size due to choice of:	500,000+	< 30,000
Exposures	Cheap-to-collect/store measures – e.g. genetic	Expensive, balanced mix of environmental and genetic measures
Outcomes	Cheap-to-collect administrative data – e.g. hospitalizations for diagnoses/deaths (dichotomous) → ↑ **SS.**	Expensive, directly measured bi-chemical physiologic, imaging, functional outcomes (often continuous) → ↓ **SS.**
Leading "Exposure-Measure Bias"	Large environmental exposure error >> genetic factor errors	"Better balanced errors" for environmental versus genetic factors
Leading to:	**Biased main effects and interaction results**	**Less biased results**

p 11

PKU ~~person~~ ~~PKU~~ condition ~~diagnosis~~

low IQs & other deficits — ~~teratogenic~~ ~~metabolic~~ teratogenic PKU fetus

improvement — decrease

diagnosis + care — neo natal — diet — hard to maintain — improvement

social support — Mandated testing — subsidy + drug classification (till 1972) — religious + ethnic identity — medical, social, psych + rehab services

wider social context — Bio-medical vs psychological — curing + care of chronic

0 — 5/6 — teenager — young adulthood