Equivocal Notions of Accuracy and Genetic Screening of the General Population

Kelly C. Smith
Clemson University, kcs@clemson.edu

Follow this and additional works at: https://tigerprints.clemson.edu/phil_pubs

Part of the Philosophy Commons

Recommended Citation
Please use publisher's recommended citation.

This Article is brought to you for free and open access by the Philosophy & Religion at TigerPrints. It has been accepted for inclusion in Publications by an authorized administrator of TigerPrints. For more information, please contact kokeefe@clemson.edu.
Equivocal Notions of Accuracy and Genetic Screening of the General Population*

Kelly C. Smith, Ph.D.

Address correspondence to:
Kelly C. Smith
Department of Philosophy and Religion
Clemson University
101 Hardin Hall
Box 341508
Clemson, SC  29634-1508
Kcs@clemson.edu

November 11, 1997

* Based on a paper delivered at the Twelfth New York Regional Conference in Medical Ethics, “Issues in Medical Ethics: 1997” -- sponsored by The Mount Sinai School of Medicine, CUNY; and the Ph.D. Program in Philosophy, The Graduate School, CUNY.
Key Words: Genetic testing; genetic screening; genetic determinism; reproductive technology; abortion; cystic fibrosis; BRCA1.

ABSTRACT

The explosive growth in genetic technology will quickly make possible an unprecedented number of tests for genetically based conditions. A necessary condition for the use of such tests without risk of harm to the patient is that they are "accurate". However, most discussions of test accuracy in the literature have equivocated between two importantly different meanings of the word. In particular, it must be kept in mind that a high analytical accuracy does not imply a high diagnostic accuracy. Questions about the diagnostic accuracy of genetic tests loom large at present given our limited knowledge of the complex etiology of disease and the distribution within the general population of the causal factors involved. Our current inability to supply patients with accurate diagnosis based on genetic information, however, is less problematic when examined in the context of new reproductive technologies such as in vitro fertilization (IVF) and intracellular sperm injection (ICSI).
"For every complex problem, there is a simple, easy to understand, incorrect answer."

-- Albert Szent-Gyorgyi

I. INTRODUCTORY COMMENTS

The growth in genetic testing technology is nothing short of explosive. The Human Genome Project, which aims to sequence all 100,000 human genes, is currently ahead of schedule and should be complete by the year 2001 -- some even predict that we will have the sequence for all genes in all living organisms by 2015 if sequencing efforts continue to grow at the present rate (1). Within the next 10 years, as many as 400 new genetic screens are likely to become available, aided by development of integrated circuit chips capable of performing thousands of hybridization probes on a single minute sample of DNA. Already, more than 4 million blood samples from newborn heel sticks are screened each year for various genetic disorders (2). Genetic screening is thus likely to be increasingly viewed as a common and accepted aspect of modern medicine, based in large part on the perception that offering patients screening is itself a harmless enterprise, and may even be required as part of the standard of care. However, offering genetic screening is often only harmless to the extent that it provides an accurate basis for making informed decisions about one's condition.
II. INFORMATION AND ACCURACY OF GENERALIZED GENETIC SCREENING

I do not mean to suggest that genetic screening is inherently evil, much less that it is avoidable. It is also beyond the scope of the present work to address all the various social and ethical issues that the widespread use of genetic screening in the general population will surely raise: access to insurance, privacy, social stigmatization, eugenics, etc. However, I do want to draw attention to a problem with genetic screening when it is applied to the general population which has not received as much attention in the literature as it deserves.

The problem is that discussions of the accuracy of genetic testing often equivocate between two distinct meanings of "accuracy". With our current state of knowledge, a high degree of analytic accuracy in a genetic test does not imply a similarly high degree of diagnostic accuracy. By "analytic accuracy" I mean the accuracy with which a test identifies the presence of the gene it is designed to locate -- avoiding both positive results when the gene is absent and negative results when the gene is present. By "diagnostic accuracy" I mean the accuracy with which the test can be used to make predictions about a patient's prognosis. For the sake of argument, let's assume that we are only discussing tests which have a 100% analytic accuracy. That is, we have somehow managed to avoid the problems posed by laboratory error and
we are testing directly for the gene in question, not a closely linked marker. Even in this ideal case, interpreting the results of the test in a way that is meaningful to patients will be extremely difficult for a long while to come.

At its most basic, the reason for this can be stated as follows: the presence of a particular gene associated with a medical condition does not insure that one will develop the condition. There are three possible explanations for this lack of "genetic determinism": multiple mutations (alleles) for a single gene, multiple genes, and gene-environment interactions. If the genetic basis of a condition has many alleles and their frequencies vary widely between populations, then the type of test required and the way the results are interpreted are relative to the population from which the patient comes in complex ways. If the condition has multiple genes, then we will need to determine all or most of the genes involved and test for each of them simultaneously to obtain an accurate diagnosis. Finally, if the condition exhibits significant gene-environment interaction -- as is the case with most complex medical conditions and particularly those involving "propensity genes" (e.g., cancer, coronary disease, etc.) -- we must be able to assess the variation of significant environmental factors within the patient population to be screened in order to make diagnostic predictions.
It is important to keep in mind that this is more than the obvious point that, “one can never predict things with absolute certainty.” With current techniques, it is very difficult even to assign an accurate probability to a particular diagnosis, though this seems a minimum requirement for the justification of testing in the first place. One way to illustrate the problem is by looking at the procedure most often initially used to isolate a gene associated with a medical condition. Since such conditions are typically quite rare, the first step is to locate a population in which the condition is unusually common so that we have more of a phenomenon to study. Then, researchers conduct statistical analyses to see if there are any regions of the chromosomes which seem to covary significantly with the condition -- that is, if there are any pieces of genetic material one is significantly more likely to have if one has the condition than if one is healthy. Once we have located a segment of DNA that appears to covary with the condition, we can use molecular techniques to refine the analysis. In the ideal case, these efforts will isolate a single gene which is then designated as “the gene for” the condition. At this point, we can develop genetic tests to determine if individuals have the gene, but can we make accurate diagnostic predictions about the likelihood of their developing the condition?

Suppose, for example, I find that in populations of Eastern European (Ashkenazi) Jews, there is a very high incidence of breast cancer
associated with a particular gene (BRCA1). Fully 85% of Ashkenazi women with BRCA1 will develop breast or ovarian cancer in their lifetimes. Now suppose I produce a test for BRCA1 and administer it to women in the general population. If a woman who is neither Jewish nor European tests positive for BRCA1, do I tell her that she has an 85% chance of contracting cancer? Such a prediction is justifiable only if one of two assumptions is correct:

1) There are no other factors besides the BRCA1 gene which significantly alter the likelihood of developing cancer.
2) There are other factors which significantly alter the likelihood of developing cancer, but these have the same distribution in the general population as they do in the Ashkenazi population and therefore do not differentially affect the likelihood of developing the condition from one population to the other.

The first assumption is almost certainly false for a condition like cancer which probably involves multiple genes with multiple alleles and complex gene-environment interactions all at once. So, is the second assumption true? If we are honest, we should admit that we really don’t know; both because we do not know all the causal factors which influence cancer and because, even if we did know them all, we have not measured their distribution within the general population. Moreover, the population in which we conducted the initial study was attractive precisely because it
was NOT representative -- in particular, it had an unusually high incidence of the rare disorder and associated gene(s). But the same factors which produced this biased distribution in the study population (e.g., inbreeding) are almost certain to have skewed distributions of other factors as well. So the study population is unlikely to have a “normal” distribution of whatever additional factors may be causally influencing the condition -- whether they be genetic, cytoplasmic or environmental. The diagnostic accuracy of the test with respect to the general population is entirely dependent on the questionable assumption of a uniform distribution of causal factors and is thus suspect -- despite the fact that we have (ex hypothesi) a test with perfect analytical accuracy.

III. CYSTIC FIBROSIS AS A CASE STUDY

Cystic Fibrosis (CF) makes a good case study for the type of problem I want to highlight for several reasons. For one thing, it has a relatively simple etiology as genetic conditions go in that there seems to be just a single gene involved. For another, it is relatively common -- about 1 out of every 2,500 whites of European descent is affected. Lastly, it is one of the few genetic conditions for which there is data on the population distribution of the alleles involved (as well as the effectiveness of education programs).
CF is a condition in which thickened secretions in the respiratory, digestive and reproductive systems result in chronic respiratory infections, loss of respiratory function, digestive difficulties and infertility. It can be an extremely debilitating condition and is often fatal, though the median survival has now climbed to forty years old and approximately one half of CF patients survive into their fifties (with intensive therapy). It is caused by various recessive autosomal mutations in the gene which codes for the Cystic Fibrosis Transmembrane Conductance Regulator (CFTR) protein, of which there are known to be more than 300 alleles\textsuperscript{11}. Since the majority of mutant alleles are carried by people who are heterozygous (and thus don’t exhibit the recessive disease), as many as one in every thirty U.S. citizens (3% of the general population) may carry at least a single copy of a CF mutation.

CF can be diagnosed in three different ways: through an assessment of symptomology/physiologic performance, through the sweat chloride test, and through genetic testing for a known mutant allele. Since fully 85% of CF cases occur in families with no known history of the condition, it has been proposed as a desirable candidate for generalized screening of the population (12, but see also 13).

As Table 1 shows, however, just about every combination of diagnostic indicators has been documented\textsuperscript{12}. There is also no simple relationship between particular genetic alleles and the severity of the disease (11, 15,
Thus, on the basis of genetic tests and/or sweat chloride values, it is very difficult to predict whether a patient will develop the condition and how severely he will be affected. Lastly, there are large differences between populations with respect to the relative frequency of the alleles and quite different screens are often required for different populations. Even for what most people intuitively consider a homogenous population, white Europeans, genetic diversity is the rule rather than the exception. For example, screening for just 4 alleles in Britain is sufficient to detect 85% of CF cases -- but in Southern Europe, screening for as many as 60 mutations detects only 75% of the CF cases.

Therefore, even for a “single gene” condition like CF, it is not enough simply to identify all the alleles which are associated with the condition. We must also assess their distribution in various populations and reliably determine into which population a patient should be classed. All three tasks present complex problems which require a great deal more work to solve adequately.

IV. THE PSYCHOLOGY OF INFORMATION

Above and beyond the information problems discussed above, there are additional problems with the way the genetic results will be perceived. There is a strong psychological tendency to view numerical information as being highly accurate, despite the fact that the actual
information content of such numbers can range from 0 (completely uninformative) to 1 (perfect predictor). In the case of genetic screening, the information content of numerical diagnostic predictions is typically unknown or, at the very least, subject to debate. To be sure, this is not a problem limited to the lay public as even professional logicians and statisticians often make elementary errors in interpreting information -- particularly statistical information (19).

However, in the context of genetic screening, this problem looms particularly large. We wish patients to be actively involved with their own treatment and require informed consent for any medical procedure done to a patient that might inflict harm. The ideal case is for the patient to assess the risks and benefits for herself and decide whether to undergo the procedure. To what extent can this ideal be met in the case of genetic screening?

The first question is, “Can the patient be made to understand the relevant details of a highly technical procedure like genetic screening?” The evidence suggests that, with a great deal of effort by trained genetic counselors, they can (20). This is encouraging, though it must be kept in mind that there is currently a severe shortage of trained counselors. Thus, if screening becomes commonplace, it will be increasingly performed by general practitioners whose training in genetics and probability is not of uniformly high quality16.
The next question is, “Are people able to retain the information accurately beyond the immediate context of testing?” Here, the evidence is less encouraging. In one three year follow-up to large-scale CF screening, it was found that 20% of those identified as carriers incorrectly recalled the results of the testing. Moreover, of those who correctly recalled the result, 46% interpreted it incorrectly as meaning that they were only likely to be carriers (20). It has been suggested that this problem may ultimately be correctable by changes in the educational system to produce graduates who are better informed about genetics and statistics\(^1\)\(^7\), but this is certainly a very long term project.

The last question is, “Do people actually want the tests?” The answer to this is complex. Offers of free genetic screening to members of the general population typically result in very low response rates, suggesting that many people have little desire to know their genetic status (23). On the other hand, it’s been found that fully 70% of those offered screening in person by a health professional accept (6). Whether a patient accepts a test depends strongly on the mode of presentation\(^1\)\(^8\), and this brings up the issue of whether even the offer of genetic testing by a physician is, in fact, a neutral act. It seems likely that such an offer will be perceived by the patient as constituting at least an indication of safety and accuracy, and perhaps even an endorsement of value. If so, this will significantly
strengthen the perception that the resulting diagnostic prediction is highly accurate.

V. GENETIC SCREENING AND REPRODUCTIVE TECHNOLOGY

A. Pre-natal Screening

One of the most obvious and compelling but also most controversial applications of genetic screening is its use in pre-natal care. It is compelling because we are deeply concerned with the health of our children -- often more so than with our own (as any pediatrician can testify). It is controversial because, although there is always the hope that early detection will allow early intervention, this is currently quite rare and thus the main outcome of a “bad” genetic result is termination of the pregnancy.

Two techniques of obtaining tissue samples for genetic analysis are currently widely used. Amniocentesis is performed at 14-16 weeks of gestation and involves needle aspiration of amniotic fluid containing fetal cells. The cells are then cultured and their DNA analyzed. Chorionic Villus Sampling (CVS) is performed at 9-12 weeks gestation and involves collection of cells from the chorion (either by catheter or needle), which can then be immediately analyzed. CVS is a major improvement over amniocentesis in that the procedure can be performed much earlier and the results obtained more quickly (24 hours as opposed to 10-14 days).
Both procedures are highly invasive and carry small but significant risk to the fetus (approximately 1% risk of fetal loss). Moreover, some conditions can not currently be detected at eight weeks (e.g., some neural tube defects). For these reasons, it is unlikely that genetic screening using these tissue collection techniques will ever become standard procedure for the vast majority of pregnancies. However, this is all likely to change soon as new techniques for isolating fetal cells from maternal blood samples come on line (24). It is now possible to detect and isolate fetal cells in the maternal blood as early as 9 weeks gestation.

As little as fifteen years ago, it would have been impossible to perform genetic analysis on the vanishingly small amounts of DNA obtained from such a tiny sample of cells, but the newly-developed Polymerase Chain Reaction (PCR) technique is capable of amplifying DNA from a single cell quickly and accurately. Moreover, as the Human Genome Project reaches completion and more and more genes are identified, it will be possible to screen very early fetal cells for their genetic contents. What this means in practice is that soon there will be techniques which are available very early in pregnancy and are no more invasive or dangerous than a routine blood sample. Under these circumstances, it seems very likely that they will quickly become an accepted part of standard care.

B. Pre-implantation Screening
With In Vitro Fertilization (IVF), it is now possible to test extremely early embryos for genetic conditions. In IVF, multiple donated eggs are collected and mixed with donated semen in vitro. It is possible to remove a single cell at the 4 or 8 cell stage of the resulting embryo and, using PCR, amplify and analyze its genes. In fact, it is possible to detect some genetic abnormalities in a single-celled zygote using Fluorescent In Situ Hybridization (FISH). FISH involves treating the cell with fluorescent genetic probes which will hybridize to some kinds of genetic abnormalities, making the deformity visible under a microscope.

Although early detection is a major advantage, IVF is also extremely invasive. The costs, both emotional and economic, are quite high -- with success rates of approximately 14% per procedure and costs of $15,000 (2). IVF pre-implantation screening will thus remain a specialized procedure for the present.

C. Pre-fertilization Screening

Techniques for the screening of individual gametes are currently under development. If techniques like FISH eventually make it possible to certify samples of sperm or eggs as “defect free”, this would constitute the ultimate in screening. For one thing, it offers advantages over screening of the parents since, even if both parents are carriers of a recessive condition, there is only a 25% chance the child will exhibit the condition.
(and a 50% chance the child will also be a carrier). For another, it avoids the issue of termination of pregnancy since undesirable pregnancies will not even be started. At the very least, the technique is likely to become popular with sperm donated to sperm banks as the collection methods are simple and artificial insemination is less invasive, cheaper and more effective than IVF23.

VI. SOME TENTATIVE CONCLUSIONS

Healthcare decisions in this country seem all too often to result from an ad hoc synergy of special interests, without adequate consideration of what is in the best interests of either the individual or the community24. In particular, it is often argued that, given the public interest in testing and the enormous amount of money to be made from generalized screening, the advent of routine genetic screens is simply a matter of time (27). Despite calls from several professional bodies that genetic screening should be preceded by careful pilot studies (2), the funding for these studies is difficult to find. The only reason the CF follow-up studies were commissioned was at the behest of the ELSI committee of the HGP, itself a highly unusual and perhaps temporary entity.

To summarize, I have tried to establish the following points which must be considered when discussing genetic screening of the general population:
1) Currently, the data we need to assess the diagnostic accuracy for most genetic conditions is simply not available. Moreover, it is difficult to secure the funding needed for such studies.

2) Members of the general public will nevertheless perceive testing as diagnostically accurate, particularly if tests are offered as a routine part of care by their physician. This will, in turn, drive the development of more and more genetic tests by private companies interested in the enormous profits to be had from screening of the general public.

3) It is very difficult (though not impossible) to educate people about the implications of genetic results, especially given the current dearth of trained counselors. It is even more difficult to insure that they retain the information for the extended periods of time often necessary to make appropriate lifestyle or treatment decisions.

So, what can we conclude about genetic screening of the general population? I would like to offer some tentative conclusions based on the two notions of accuracy I have distinguished:

1) Individual patients are simply not in a position to evaluate the
diagnostic import of genetic tests. As individuals, they can not assess population level phenomena and as non-experts, they can not be expected to follow all the technical minutia. The situation seems analogous to the public release of new drugs: patients might be interested in trying a new drug but we typically do not allow this until we have had a chance to carefully assess its effectiveness and possible side effects. Genetic screening of the general population of adults should fall in the same category. To be sure, there will be special situations -- for people in high risk populations, we have more data on the distribution of causal factors and thus a better grasp of the diagnostic accuracy. Moreover, a total ban on screening would prevent the collection of precisely the data we need to evaluate its effectiveness. However, to the extent that we allow screening of members of the population not known to be at high risk, it should be for the purpose of data collection only and they should be treated as any other experimental subject.

2) In cases of pre-natal screening, the same information problems apply as with the general adult population. However, to the extent that early intervention may be possible, experimental evaluation of early treatments for genetic conditions is certainly justified. Moreover, allowing limited screening of fetuses may be a good
way to generate the data needed to assess the diagnostic accuracy of the screens. Of course, this would require extensive follow-up after the initial screen to determine which children went on to develop the condition and which did not.

3) Pre-implantation and pre-fertilization screening pose much less severe problems with respect to diagnostic inaccuracy than those associated with other screens of the general population. This is because a selection often must be made as to which embryo(s) to implant or which gamete(s) to employ. If a decision such as this can not be avoided, and if there is no alternative means of selection which is more informative, it is less problematic to rely on information of uncertain diagnostic accuracy provided by current genetic screening techniques. This is because the decision perse does not have to be justified on the basis of the information, since a decision of some sort is unavoidable. Of course, the precise form of the decision is justified via the genetic information, but the alternative is often to use either arbitrary or morally questionable criteria (e.g., sex selection). Therefore, even if the genetic screening turns out to be completely uninformative (which is unlikely), it is hard to see how we are any worse off using it in these cases.
<table>
<thead>
<tr>
<th>Classic Symptoms</th>
<th>Sweat Test</th>
<th>Genetic Test</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>Positive</td>
<td>Positive</td>
<td>Classic CF</td>
</tr>
<tr>
<td>Positive</td>
<td>Positive</td>
<td>Negative</td>
<td>Occurs</td>
</tr>
<tr>
<td>Positive</td>
<td>Negative</td>
<td>Positive</td>
<td>Common</td>
</tr>
<tr>
<td>Positive</td>
<td>Negative</td>
<td>Negative</td>
<td>Common</td>
</tr>
<tr>
<td>Negative</td>
<td>Positive</td>
<td>Positive</td>
<td>Occurs</td>
</tr>
<tr>
<td>Negative</td>
<td>Positive</td>
<td>Negative</td>
<td>Occurs</td>
</tr>
<tr>
<td>Negative</td>
<td>Negative</td>
<td>Positive</td>
<td>?</td>
</tr>
<tr>
<td>Negative</td>
<td>Negative</td>
<td>Negative</td>
<td>Health</td>
</tr>
</tbody>
</table>
REFERENCES


ENDNOTES

1 For example, The American College of Obstetrics and Gynecology includes the offering of prenatal screening as part of the standard of care (3). For a more detailed discussion of the evolution of standard of care as it relates to testing, see 4, 5.
For example, a bilateral radical mastectomy may (or may not) be a justified procedure if one is at high risk for breast cancer. However, such a procedure obviously becomes more unwarranted as the prognosis becomes less certain.

An exception to this is the work of Benjamin Wilfrond (6, 7).

This problem is, in some sense, a transient one since it will be corrected as our knowledge of the etiology of disease grows. However, this makes it no less serious an issue. Moreover, the information required to resolve the problem will be far more difficult to collect than many currently realize.

It should be noted, however, that a small amount of analytical error can seriously erode the diagnostic accuracy of a test for rare conditions. For example, even with a relatively common genetic disease such as Cystic Fibrosis and a test with a 99% analytical accuracy, a positive result is twenty times more likely to be due to test inaccuracy than the presence of the gene.

For a related discussion concerning the causal nature of disease in general and phenylketonuria in particular, see (8, 9)
Whatever else it might do, widespread use of new cloning technologies on humans would be an interesting test of the extent to which human traits are the result of genes rather than environments. By holding the genetic background constant and varying both the cellular and maternal environment, variation in traits between donor and clone would be attributable to differences in the environment -- at least more so than they are at present.

I do not mean, of course, that it is difficult to calculate a probability, but that it is difficult to defend such numbers as being justified, given our current state of knowledge about the etiology of most genetic conditions.

Of course, the research does not stop here. Next we would want to know (at least) the amino acid sequence of the protein, as well as its three-dimensional confirmation, location and function within the cell.

The exact incidence figures cited varies from 1/4700 to 1/1000 because different authors examine subtly different populations.

CFTR is a cAMP mediated chloride channel protein found in the plasma membrane. Defects in this protein result in poor uptake of chloride ions by the cells, which results in high levels of chloride in the sweat (10). Thus, the
sweat chloride test is currently considered the gold standard for CF diagnosis. It should be noted that a single mutation, delta F508, seems to account for a large majority of CF cases in whites (11).

12 This raises interesting questions about the nature of a disease - should it be defined in terms of a symptomology, as was the case in earlier times; or should identification of putative causal agents be considered primary (see 14)?

13 see (16, 17). Note also that these are known cases of CF - it may be that some cases of CF are not detected (due to very mild symptoms, etc.) and thus are not reflected in these numbers.

14 This is not a simple matter of patient self-identification. As anyone who has done family history research knows, people rarely have very clear ideas about their ancestry more than two or three generations back. What is worse, they often have incorrect ideas.

15 New techniques may be able to avoid some of these difficulties, at least for single gene conditions, by a direct assay of protein function. It has been found that many cells in easily accessible tissue (white blood cells, for example) express minute amounts of mRNA from genes that they do
not “officially” express. These cells can be isolated from a blood sample, the mRNA amplified and translated into protein, and the protein activity directly assayed (18). Therefore, at least to some extent, it is not necessary for us to know precisely which mutation is involved or even its precise effect.

16 In fact, given the performance of many physicians on other genetic tests like the APC test for familial adenomatus polyposis, the ability of the average U.S. physician to act as a genetic counselor without specialized training is in serious question (21).

17 It has even been argued that, on certain conceptions of autonomy, patients have an obligation to thoroughly inform themselves about testing (22).

18 Women often report that they find it difficult to refuse screening which is offered as a routine part of prenatal care -- as with alpha fetaprotein tests (4).

19 For example, there is no conclusive evidence that CF screening enables effective early intervention, despite much early optimism (7).
There is the difficulty that cells from the fetuses of earlier pregnancies will sometimes still be in the maternal bloodstream (25). This poses no problem, of course, for first pregnancies and future refinements may allow us to differentiate between the two cell lines before performing genetic analysis.

At present, this technique can only be used to detect relatively gross abnormalities such as major chromosomal rearrangements or aneuploidies.

New techniques in embryo collection involve using intrauterine lavage to remove an embryo at the blastocyst stage (26). If these sorts of procedures become routine, pre-implantation screening will as well.

Intracytoplasmic Sperm Injection (ICSI) even makes screening of individual sperm before fertilization a possibility.

Wilfrond labels this the “extemporaneous model” of decision making and advocates a more “evidentiary model” instead (6, 7).