

## SCIENCE AND SOCIETY

# Challenges of translating genetic tests into clinical and public health practice

Wolf H. Rogowski, Scott D. Grosse and Muin J. Khoury

**Abstract** | Research in genetics and genomics has led to an expanding list of molecular genetic tests, which are increasingly entering health care systems. However, the evidence surrounding the benefits and harms of these tests is frequently weak. Here we present the main challenges to the successful translation of new research findings about genotype–phenotype associations into clinical practice. We discuss the means to achieve an accelerated translation research agenda that is conducted in a reasonable, fair and efficient manner.

Molecular genetic testing has an increasingly wide range of applications in health care. The use of molecular genetic tests (hereafter termed ‘genetic tests’) for diagnosis and for screening within families has increased substantially. For example, the diagnosis of inherited cancers, such as Lynch syndrome, is increasingly based on molecular testing rather than family history and phenotype<sup>1,2</sup> (BOX 1). Molecular genetic testing can detect both highly penetrant gene mutations (for example, *BRCA1* or *BRCA2* mutations) and polymorphisms (for example, *APOE4* variants). In fact, the use of information on polymorphisms that are associated with moderately increased risk of common diseases such as diabetes or cardiovascular disease might at some point complement or substitute the risk calculation based on family history<sup>3</sup>. Testing for variants that can alter drug metabolism and that therefore potentially affect drug safety and efficacy is also growing, although it remains controversial<sup>4</sup>.

The costs of many new genetic tests outweigh the potential savings from early prevention, and contribute to a net increase in health care expenditure<sup>3</sup>. The costs of meeting the demand for all genetic tests could seriously strain the health care

budgets of many countries. A recent study showed that a comprehensive cascade screening programme for single-gene disorders would involve testing a substantial proportion of the population<sup>5</sup>. Funding the rising number of predictive genetic tests<sup>6</sup> would add further burden to tight budgets. Hence, third-party payers must consider carefully which genetic tests to include among the services they cover.

“...if the required levels of evidence are undefined or excessively low, this could prematurely move technologies towards practice, with potentially adverse consequences for patients and the health care system.”

The evidence surrounding the benefits of new genetic tests is frequently weak. Decision makers face a dilemma: if only tests with a high level of evidence of clinical utility are accepted, genetic and genomic discoveries could be kept from clinical application for years, with large disincentives for further research and development.

Conversely, if the required levels of evidence are undefined or excessively low, this could prematurely move technologies towards practice, with potentially adverse consequences for patients and the health care system<sup>7</sup>.

This paper addresses this ‘evidence dilemma’ in genomic medicine. Recently, health economic methods such as value of information analysis (VOI analysis) have become available to inform decisions about whether to conduct further research into a new technology.

Decision analytic methods, such as VOI analysis, can be useful in prioritizing genetic tests and further research, especially when the impact on health and budgets is potentially high. However, they cannot substitute for ethical analysis and prudent judgment. Furthermore, VOI analysis can only address selected types of uncertainty<sup>8</sup>.

The following sections provide an overview of the issues that need to be considered when balancing the benefits and harms of genetic tests, and when deciding whether to adopt a genetic test or to carry out further translational research. We conclude by recommending how decision makers can address three types of challenges: assessing evidence; ensuring that this evidence is used in decision making; and ensuring that valid prioritization tools are put in place before a test is incorporated into health care systems.

### Assessing benefits and harms

**Medical benefits and harms.** The benefits and harms of both genetic and non-genetic tests are assessed in a similar way<sup>9</sup>. The main differences lie in specific difficulties of collecting appropriate data and in the ethical implications for family members of genomic knowledge.

Assessing the benefit of genetic tests can be subdivided into three stages, which involve measuring analytic validity, clinical validity and clinical utility<sup>10</sup>. Although analytic validity is typically high for genetic tests, clinical validity is frequently less straightforward to demonstrate because of low or unknown disease penetrance. Furthermore, population diversity and bias can lead to an overestimation of the disease susceptibility conferred by a genetic polymorphism<sup>11</sup>.

It is even more important to judge whether, and how, genetic information shows clinical utility — that is, whether it guides clinical management and improves health outcomes by reducing morbidity or mortality<sup>12,13</sup>. The magnitude of medical benefit depends on the availability and effectiveness of preventive or curative therapies (BOX 1) or on the ability to avoid potentially harmful interventions in the case of negative test results. For example, at-risk children who test negative for retinoblastoma would avoid the potential harms caused by general anaesthesia.

Psychological and behavioural issues can be important in the establishment of medical benefit. Not all individuals fully adhere

to the method of prevention (for example, increased medical surveillance), even in the case of severe diseases such as hereditary cancer<sup>14</sup>. For conditions with low penetrance, adherence to preventive strategies may be lower<sup>15,16</sup>.

Genetic testing can also lead to medical harm. For example, although a positive result of a genetic test for hereditary breast cancer may improve health by taking measures to prevent the condition, it can also increase anxiety and distress. However, it was shown recently that the adverse psychological outcomes of testing positive for *BRCA1* or *BRCA2* mutations seem to be temporary<sup>14,17,18</sup>.

**Non-medical benefits and harms.** Many genetic tests are designed to detect mutations that cause highly penetrant, single-gene disorders that lack a treatment strategy. The primary purpose of these tests is to provide non-medical benefit in terms of prognostic information; for example, to inform reproductive or other life decisions. For instance, testing for Huntington's disease cannot alter the course of this lethal condition but does allow carriers and non-carriers to prepare for the future.

Genetic testing must involve informed consent, but the difficulties in ensuring that this requirement is met can lead to harm. If genetic information is shared in a family, the right of an individual to decline receiving information may be infringed<sup>19</sup>. If, for example, a young woman conducts a DNA test for a *BRCA1* or *BRCA2* mutation that has been found in her grandmother, a positive test result implies that her mother is positive for the mutation as well, even if she declined to undergo the genetic test. Informed consent is also difficult to obtain if parents make decisions on behalf of their children. Qualitative studies have reported difficulties that are inherent in ensuring that patients consent in a truly informed and voluntary manner<sup>20</sup>. Also, patients can easily misunderstand the results of genetic tests, for example, with respect to information on relative risk<sup>21</sup>.

Other potential non-medical harms of genetic testing include potential stigmatization or unfair discrimination in the labour or insurance markets. Several countries have established laws that address these issues<sup>22</sup>.

**Financial benefits and costs.** Genetic testing can incur financial benefits if it prevents costly medical conditions and complications, or if it avoids disease surveillance in mutation-negative relatives. However, it also consumes financial resources: the cost of collecting specimens, laboratory testing, counselling, follow-up testing, treatment and prevention (including for false positives), and, in the case of screening programmes, the costs of ensuring test uptake<sup>3,23,24</sup>.

In most screening programmes, the direct costs of screening typically outweigh the savings in medical costs among first-degree relatives<sup>3</sup> (BOX 1); exceptions do exist, however, such as in genetic screening for familial adenomatous polyposis<sup>25</sup> or for retinoblastoma<sup>26</sup>. This expense is not mitigated by the rapid fall in the laboratory costs of DNA tests: in fact, the total costs associated with genetic testing might increase over time because of increasing options for follow-up testing and treatment.

## Box 1 | Assessing cascade genetic testing: Lynch syndrome

Colorectal cancer (CRC) is a common disease responsible for an estimated 52,000 deaths per year in the United States<sup>2,10</sup> and 146,000 per year in the European Union<sup>54</sup>.

Approximately 3% of newly diagnosed cases of CRC can be attributed to Lynch syndrome. This autosomal dominant condition is caused by mutations in one of several mismatch repair genes, which can reliably be identified with existing laboratory tests. Relatives who inherit a mutation have a ~45% risk of developing CRC by the age of 70. Genetic testing to detect Lynch syndrome in individuals with newly diagnosed CRC has therefore been proposed as a strategy to reduce CRC morbidity and mortality in their relatives. As we explain below, these decisions are supported by evidence of the clinical utility of the intervention, and of its cost-effectiveness.

### EGAPP review of effectiveness

The Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Working Group (EWG) is a multidisciplinary panel of experts that assesses the impact of new genetic technologies (BOX 3). It constructed a chain of evidence that linked genetic testing for Lynch syndrome in patients with newly diagnosed CRC with improved health outcomes in their relatives<sup>1,2</sup>. Adequate evidence was found for testing uptake rates, adherence to recommended surveillance activities, number of relatives who could be found and that consented to testing, acceptably low harms associated with additional follow-up, and effectiveness of routine colonoscopy in reducing disease incidence.

The EWG found sufficient evidence to recommend offering genetic testing for Lynch syndrome to reduce morbidity and mortality from CRC, but insufficient evidence to recommend a specific genetic testing strategy among the several strategies examined.

### Cost-effectiveness analysis

Cascade screening for Lynch syndrome involves two steps: first, detecting the causative mutation in the proband with CRC and, second, testing whether family members have inherited the same mutation. As full gene sequencing and deletion analysis is expensive, CRC patients typically undergo pre-selection on the basis of age, family history and other phenotype-based criteria. Different testing strategies for Lynch syndrome among CRC patients have been assessed by published health economic evaluations, with cost-effectiveness ratios ranging from approximately US\$2,000 to \$12,000 per life year gained (LYG) for the most favourable strategies<sup>3</sup>. The type of testing strategy greatly influenced the economic results. For example, the incremental cost-effectiveness ratio in one study ranged from \$11,865 per LYG to \$267,548 per LYG<sup>55</sup> depending on whether patients were selected for gene sequencing according to established guidelines (the Bethesda guidelines) or whether gene sequencing occurred without selection, respectively. A new but preliminary cost-effectiveness analysis based on the EGAPP review suggests that testing all newly diagnosed CRC patients using a preliminary immunohistochemistry test is cost-effective (S.D.G., unpublished observations).

Economic evaluation can thus help to identify testing algorithms and target groups that maximize the value of the resources spent on screening and prevention<sup>4,23</sup>. For example, a national programme in the Netherlands that, since 1994, has promoted cascade screening of relatives of patients with familial hypercholesterolaemia has been shown to be cost-effective, and the National Health Service in England recently decided to adopt a similar programme<sup>56</sup>. Frequently, country-specific analyses are necessary to account for the large differences of treatment patterns and costs across health care systems<sup>57</sup>.

**Challenges in assessing benefits and harms.**

As mentioned above, it is difficult to assess the full cost of genetic testing (BOX 2); however, assessment of its benefits is also challenging. For highly penetrant monogenic diseases, mutation prevalence is low and so studies with large sample sizes are needed to robustly establish a health risk. It is frequently neither ethically nor financially feasible to assess preventive interventions in randomized clinical trials. Instead, the assessment of benefit frequently must rely on evidence relating to separate components of a programme (for example, evidence relating to the diagnostic accuracy of predictive testing and evidence on the impact of early treatment on health)<sup>22,23</sup>, and decision analytic models may be needed to quantify the magnitude of medical benefits and costs<sup>16</sup>.

The challenges of assessing benefits and harms depend on the aim of the genetic tests. In genetic screening, benefits frequently occur many years after taking the test. Therefore, the predicted health effects can be overestimated because adherence to prevention may fall below expectations, or treatment of disease may improve over time. Also, assessing potential non-medical harms, such as stigmatization or the infringement of patient autonomy, are particularly important because screening tests are offered to a population or subpopulation rather than to individuals who have actively sought clinical care or advice<sup>21</sup>. For tests that aim only to provide information, the size of the benefit to the tested person is difficult to estimate and to compare with test costs. In these cases, willingness to pay, rather than medical endpoints, may be a better indicator of benefit<sup>27</sup>.

The need for evidence is also crucial for highly penetrant conditions because the severity of disease can easily be overestimated, as in the case of hereditary haemochromatosis<sup>16</sup> or cystic fibrosis<sup>28</sup>. In other cases, there may be groups of patients or disease manifestations for which the evidence of benefit from prevention is weak. For example, female carriers of a Lynch syndrome mutation have the same risk of developing endometrial cancer as of developing colorectal cancer, but there is insufficient evidence that surveillance for endometrial cancer will save lives<sup>1</sup>.

**Decision making and scarce resources**

Despite uncertainty, decisions about the allocation of scarce resources have to be made. These decisions address two aspects: whether to adopt a new technology such as a

**Box 2 | Evidence gaps in pharmacogenetic testing: the case of HER2**

Approximately 20% to 30% of breast cancer patients overexpress human epidermal growth factor receptor 2 (*HER2*); such cases can be identified with a genetic test and treated effectively with the drug trastuzumab. The test and the drug were approved by the US Food and Drug Administration in 1998 for patients suffering from metastatic breast cancer. After 2005, its use was expanded to *HER2*-positive patients with early-stage breast cancer. *HER2* testing is now typically recommended for patients with invasive breast cancer, and only patients with positive test results are recommended for trastuzumab treatment. *HER2* testing is among the most well-known and investigated examples of testing to target treatment<sup>58</sup>.

Nevertheless, a recent commentary by Phillips points out that a number of gaps in the evidence persist<sup>58</sup>. First, there is no consensus about optimal testing methods to determine *HER2* status. Guidelines recommend using either immunohistochemistry, with inconclusive results confirmed by fluorescence *in situ* hybridization (FISH), or FISH alone. Although FISH is a better predictor of response to treatment, immunohistochemistry is substantially cheaper and is easier to perform in community laboratories. Despite the clinical success of trastuzumab, the accuracy and interpretation of *HER2* tests is an issue of concern that remains to be addressed<sup>58</sup>.

In addition, little is known about whether all eligible patients are tested, which testing methods are used and whether inconclusive results are confirmed. In a study of a sample of Medicare enrollees in the United States, only 32% of women newly diagnosed with invasive breast cancer were documented in claims data as having undergone a *HER2* test. It is unclear if eligible patients did not receive *HER2* testing because of financial or clinical reasons, or if their testing was not documented<sup>58</sup>.

Furthermore, it has been reported that a substantial percentage of *HER2* tests performed by community laboratories are inaccurate<sup>59</sup>. However, it is unclear why this is the case and what the implications of inaccurate tests are for patients<sup>58</sup>. It is also unclear how many patients receive trastuzumab despite a negative or inconclusive test result, or which testing approach is most cost-effective in clinical practice<sup>58</sup>.

Efforts to address these gaps are currently underway. These include efforts to document and address gaps on testing technologies, to standardize testing procedures and interpretation, and to provide incentives for closing these gaps<sup>58</sup>.

genetic test on the basis of existing evidence, and whether to conduct further research to support future decisions<sup>29</sup>.

**Technology adoption.** Various criteria exist to determine whether a new technology should be covered by third-party payers. Typically, the criteria are stated in legal terms. For example, the US Medicare programme funds technologies that are considered “reasonable and necessary” and the German Statutory Health Insurance only funds technologies that meet the criteria of “expedience, necessity and efficiency”.

The US Medicare programme and the German Statutory Health Insurance typically interpret these criteria with respect to evidence of patient-relevant health benefit. Thus, national-level decisions typically involve systematic reviews of the scientific literature<sup>30</sup>.

However, if funding decisions were based only on the level of evidence, technologies such as genetic testing would tend not to be covered because, as discussed above, evidence of their benefit to health can be difficult to establish. In addition, assessments that address only the evidence of a technology’s benefits and not its costs may not be sufficient to address the problems

involved with escalating health care costs<sup>31</sup>. Therefore, formal decision processes in Europe, North America and Australia increasingly include the use of economic techniques, such as cost-effectiveness analysis (CEA), cost-utility analysis (CUA) and cost-benefit analysis (CBA)<sup>32</sup>. CUA and CBA integrate medical endpoints in a single indicator that can be compared with the costs involved. Economic techniques are applied in decision making procedures by the English National Institute of Health and Clinical Excellence (NICE); for example, their recommendation in 2002 to use trastuzumab in women with *HER2*-positive breast cancer (BOX 2) (documentation is provided at the NICE website; see the further information box).

Although quantitative decision analytic techniques can facilitate more transparent and evidence-based decision making, it consumes resources and it is difficult to formalize all the aspects that are relevant for priority setting. For example, quality-adjusted life years, which are used in CUA, provide one solution to the difficult problem of aggregating and comparing the different health benefits (such as effects on pain, anxiety or mobility) in one single measure. However, quality-adjusted life

**Box 3 | Translational genomics research: the EGAPP initiative**

The Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Working Group (EWG) is an independent panel established in 2005 to develop a systematic process for evidence-based assessment of genetic tests and other applications of genomic technology. Its key objectives are: to develop a transparent, publicly accountable process; to minimize conflicts of interest; to optimize existing evidence review methods to address the challenges presented by complex and rapidly emerging genomic applications; and to provide clear links between the scientific evidence and the subsequently developed recommendation statements. The EWG is composed of multidisciplinary experts in areas such as clinical practice, evidence-based medicine, genomics, public health, laboratory practice, epidemiology, economics, ethics, policy, and health technology assessment<sup>10</sup>.

EGAPP methods draw on the experiences in the *ACCE* project, which took its name from the four components of evaluation of genetic tests — analytic validity, clinical validity, clinical utility and the associated ethical, legal and social implications. As, typically, no single body of evidence can establish the benefits gained from a genetic test, the EWG aims to construct a chain of indirect evidence about a series of key questions that link the genetic test with health outcomes. The clinical evidence for these questions is graded according to acknowledged criteria. Practice recommendations for clinicians are issued, depending on the magnitude of net benefit, the certainty of evidence, and consideration of other clinical and contextual issues<sup>10</sup>.

EGAPP methods differ from those of other evidence review processes, which, for example, do not formally evaluate analytic validity. Most currently available genetic tests are offered as laboratory developed tests and are not currently reviewed by bodies such as the US Food and Drug Administration. Therefore, EGAPP reviews include an analysis of the laboratory data needed to assess analytic validity; for example, they assess the average analytic sensitivity and specificity of assays by distributing standardized specimens to multiple laboratories<sup>10</sup>.

years are not sufficiently sensitive to capture all relevant benefits from genetic testing<sup>27</sup>. Furthermore, economic evaluations of genetic tests frequently lack methodological quality and standardization<sup>3,33</sup>. As a result, estimates of the cost-effectiveness of a given intervention can vary widely, thereby making it difficult to synthesize results across published analyses<sup>3,33,34</sup>. For example, a recent economic evaluation of population screening for *HFE* mutations associated with hereditary haemochromatosis<sup>17</sup> reached very different conclusions to previous studies<sup>3,35</sup>. In addition, even if it is possible in principle to address a range of aspects, such as equity concerns<sup>36,37</sup>, CUEs typically focus on the maximization of health or welfare rather than on the wider principles of distributive justice.

**Further translational research.** Several criteria and methods can be applied to decide whether to carry out further research, although decision makers typically rely on less formalized approaches than those that are used to decide on technology adoption<sup>38</sup>. A health economic technique that is used to estimate the potential value of further research from a health care system perspective is VOI analysis, which is an extension of CEA and CUA<sup>39</sup>.

The expected value of perfect information is the difference between the net benefit a decision maker could achieve if all uncertainty was resolved and the expected net

benefit achieved if the decision is based on the current evidence. In a Bayesian framework, the value of further evidence depends on the costs of a wrong decision, which is based on existing evidence in terms of the value (for example, health) that is forgone, and the usefulness of this further evidence to avoid such error. The expected value of perfect information can be estimated for the decision and for selected parameters, so that the collection of further evidence can focus on those tests or parameters about which evidence is likely to be of highest value to the decision maker. For example, a recent VOI analysis on screening for hereditary haemochromatosis in Germany found that the parameter associated with the highest cost of uncertainty was the level of adherence to prophylactic treatment<sup>8</sup>.

However, the use of VOI analysis to prioritize further research has limitations. For example, it presumes an agreed cost-effectiveness threshold, which may be neither empirically valid nor ethically justified<sup>40</sup>. In addition, VOI analysis mainly addresses parameter uncertainty (for example, of an estimated mortality risk) rather than structural uncertainty (for example, whether all potential harms from testing are appropriately accounted for in the model) and full VOI calculations can be computationally extensive. Although decision analytic methods can account for all potential harms that are represented in the model, they typically do not account for

the precautionary principle; they treat health gains and risks equally, regardless of whether they are associated with commission or with omission.

A key challenge in translational research is to assess the clinical validity and utility of a genomic application that is ready to be put into practice (for example, a predictive or pharmacogenomic test), so that evidence-based guidelines can be developed. To support this type of research a number of initiatives have developed frameworks that harmonize the assessment of genetic tests and that account for medical, economic and ethical aspects<sup>21</sup>. An example is the US Evaluation of Genomic Applications in Practice and Prevention (*EGAPP*) initiative<sup>10</sup> (BOX 3).

The current EGAPP methods are based on a qualitative rather than a quantitative synthesis of the evidence, so they do not allow for VOI calculations. However, the systematic literature reviews prepared for EGAPP can provide a valuable basis for cost-effectiveness modelling, which can be used to establish the most cost-effective testing algorithms (BOX 1). A CEA of different testing algorithms for Lynch syndrome based on the EGAPP review is currently in progress (BOX 1).

VOI analysis can guide priority setting in a later stage of translational research; for example, for prioritizing research designs of randomized pilot studies of such screening programmes<sup>8</sup>. Furthermore, VOI analysis can provide insights for research in earlier phases. The expected VOI depends on the probability of making a wrong decision, the consequences of a wrong decision and the population affected by the decision. Indeed, the criteria for preliminary ranking of topics by the EGAPP initiative include an assessment of whether or not wrong decisions about the use of the tests are possible or probable and of how many individuals are likely to be tested<sup>10</sup>.

**Prioritizing technologies and research**

Prioritizing the funding of technology and further research involves balancing cost arguments and the evidence-based assessment of health benefit — all in the context of limited financial resources<sup>30,31</sup>.

Ideally, decisions would be based on a rigorous evidence-based assessment and appraisal of the scientific evidence; however, this is frequently not achieved<sup>41,42</sup>. If the level of evidence were the sole criterion of decision making, the costs of research might become excessive — to the extent that they might exceed the potential

value of the test to decision makers. Furthermore, this approach may favour technologies with patent-holding manufacturers, who can expect to profit from the sales<sup>31</sup> of genetic tests. VOI analysis provides helpful insights into other criteria for determining the necessary level of evidence — further evidence should be collected if the benefit it provides in terms of reducing costs associated with decision uncertainty is likely to exceed the cost of research. However, estimating the VOI consumes scarce resources and thus cannot be applied to all genetic tests.

As resources are limited, costs frequently remain a criterion in decision making. If costs are not considered in an explicit and rational manner, there is a risk of unnecessary benefit forgone to the health care system, either because services with comparatively low cost per health gain are excluded from coverage in support of less efficient services<sup>43</sup>, or because prioritization decisions are left to the contingencies of day-to-day medical practice<sup>44</sup>. Although theoretical concepts have been developed to assess the cost-effectiveness of novel health technologies<sup>29,45</sup>, they are typically not applied in decision making because of the political or legal impediments to explicitly using cost-effectiveness as a criterion for health care coverage. Also, the tools for making decisions about novel health technologies are not yet sufficiently adapted to genetic testing — decisions about technologies, such as genetic tests, that are comparatively inexpensive to health care payers (in aggregate terms) are frequently made on local rather than national levels<sup>30</sup>.

For genetic tests with uncertain clinical relevance, the result of prioritization decisions may be to exclude them from health care coverage. However, leaving the provision of genetic tests to the free market might cause harm to patients and health care systems. This is because without appropriate counselling, genetic tests may cause unnecessary fear for patients<sup>46</sup> and unnecessary costs for health care systems if concerned patients are provided with follow-up tests for medically meaningless genomic profiles<sup>24</sup>.

Prioritization decisions inevitably involve value judgments regarding the consideration of health-related need, medical benefit or equity, which may differ across countries and health care payers. Typically, not all relevant aspects can be incorporated into the methods of formal analysis. Therefore, the outcomes of economic analyses should not determine but inform decisions, which need to be made in a deliberative process. Procedural standards

such as transparency or right to appeal can help solve conflicts between value judgments, both in the development of principles and in their application<sup>47</sup>.

### From discovery to clinical practice

Inevitably, decisions that balance the aims of health care innovation and financial sustainability have to be made, and it will not be sustainable for health care payers to fund the use of all genomic discoveries or to fund rigorous translational research for all novel genetic tests and those yet to be developed.

**Priority setting for genetic tests.** To ensure fair and reasonable translation of genetic tests into clinical care, the issue of prioritization needs to be addressed explicitly by collaborative efforts of geneticists, ethicists, health economists, representatives of the targeted population and decision makers. Although the ranking of criteria may differ across health care systems, evidence of effectiveness always has an important role. Thus, evidence reviews such as those provided by EGAPP are needed<sup>10</sup>.

Probabilistic decision analytic models provide powerful means to synthesize the decision-relevant evidence on effects and costs of genetic tests in a structured manner. Decision analytic economic evaluations can provide helpful guidance, especially

for genetic services with a potentially high impact on health care budgets, such as population screening programmes<sup>16,23</sup> (BOX 1). Thus, such modelling activities are also needed. However, the results of such studies can easily be biased owing to selective incorporation of evidence and insufficient sensitivity analyses. Also, their results may need to be adapted to different circumstances across nations; therefore, studies must feature a high level of transparency. Authors and peer reviewers thus need to put more emphasis on the methodological requirements set by health economic guidelines<sup>33,48</sup>.

Transparency and participation in national decision processes can help to ensure that both scientific evidence and other decision-relevant criteria are accounted for appropriately in decision making, and that value conflicts are resolved in an accountable and reasonable manner<sup>47</sup>.

The decisions about most genetic tests are, however, likely to be made by health care providers; such decisions also warrant pragmatic criteria for prioritization. This issue is currently being addressed by *EuroGentest*, a network of excellence funded by the European Union<sup>49</sup>. Based on the needs of medical practice as well as on ethical considerations, a set of criteria for prioritizing genetic tests under the constraints of a limited budget is being

## Glossary

### Analytic validity

The ability of a genetic test to accurately and reliably measure the genotype (or analyte) of interest in the clinical laboratory, and in specimens that are representative of the population of interest.

### Cascade screening

Active screening among the family members of known carriers of disease-relevant mutations.

### Clinical validity

The ability of a genetic test to accurately and reliably predict the clinically defined disorder or phenotype of interest.

### Clinical utility

The improved measurable clinical outcomes shown by a genetic test, and the usefulness and added value to patient management resulting from a test compared with current patient management without genetic testing.

### Cost–benefit analysis

Estimation of the costs of an intervention and the value of outcomes, which can include both health and non-health outcomes, in monetary terms.

### Cost-effectiveness analysis

Estimation of the net cost per unit of health outcome gained, such as cases of disease, number of deaths prevented or life years gained.

### Cost–utility analysis

A variant of cost-effectiveness analysis in which a preference-based measure of health, such as the quality-adjusted life year, combines information on mortality and morbidity.

### Distributive justice

Normative principles designed to guide the allocation of the benefits and burdens of economic activity.

### Molecular genetic test

The application of genetics and genomics research results to the analysis of human DNA, with the aim of detecting disorders that have a genetic basis or specific reactions to medical treatments.

### Quality-adjusted life years

A summary health measure that combines information on mortality and morbidity by assessing different dimensions of health with a generic instrument and then evaluating these different health states by weights elicited from representative surveys.

### Value of information analysis

Assessment of the potential benefit of further information for a decision maker; value of information depends on the probability that the decision based on existing information is wrong and the consequences of a wrong decision, in terms of resources and health benefit forgone.

developed. The criteria are being developed in a consensus process of geneticists involved in local decision making, genetics researchers, health economists, ethicists and patient representatives.

As prioritization may lead to the exclusion of genetic tests from coverage, the regulation of direct-to-consumer marketing of genetic tests should be addressed similarly by geneticists, ethicists, lawyers and policy makers<sup>50</sup>.

**Priority setting for research.** Among numerous gaps in the evidence on genetic testing (BOX 2), translational research needs to focus on those that are most relevant. To achieve this, criteria for the prioritization of translational research need to be established<sup>7</sup>. For example, the EGAPP initiative is developing its methods for selecting promising candidates for further translational research.

VOI calculations<sup>8</sup> and potential indicators of high VOI<sup>51</sup> — as well as other criteria used in decision making about health technology assessment<sup>52</sup>, health research<sup>38</sup> and coverage with evidence development<sup>53</sup> where reinforcement for a technology is tied to participation in a research protocol — may be helpful in this process of priority setting. Potential indicators of high VOI include high test use, low evidence, high budget, high health impact and heterogeneity of use in clinical practice. Additionally, tests may be selected by factors indicating that improved evidence is likely to be used in clinical practice, for example, the controversial nature of a genetic test, and appropriate timing of the analysis or information needs stated by health care providers.

Given the potentially high impact of further translational research on public budgets, sophisticated methods for prioritization may be more necessary here than for assessing the tests themselves.

**Conclusions**

Challenges in translating genetic tests from research to practice are being addressed on a global scale. These include the prioritization of genetic tests and of further applied research. Prioritization needs to incorporate quantitative analysis, qualitative weighting of ethical and other principles, procedural fairness, and political pragmatism in the specific context of genetic tests. Prioritization of translational genetic and genomic research is as much a challenge as prioritization of the genetic tests themselves.

Wolf H. Rogowski is at the Helmholtz Center Munich, German Research Center for Environmental Health (GmbH), Ingolstädter Landstrasse 1, 85764 Neuherberg, Germany.

Scott D. Grosse is at the National Center on Birth Defects and Developmental Disabilities, Centers for Disease Control and Prevention, 1600 Clifton Rd, MS E88, Atlanta, Georgia 30333, USA.

Muin J. Khoury is at the Office of Public Health Genomics, Centers for Disease Control and Prevention, 1600 Clifton Rd, MS E61, Atlanta, Georgia 30333, USA.

Correspondence to W.H.R.  
e-mail: [rogowski@helmholtz-muenchen.de](mailto:rogowski@helmholtz-muenchen.de)

doi:10.1038/nrg2606

Published online 9 June 2009

1. Palomaki, G. E., McClain, M. R., Melillo, S., Hampel, H. L. & Thibodeau, S. N. EGAPP supplementary evidence review: DNA testing strategies aimed at reducing morbidity and mortality from Lynch syndrome. *Genet. Med.* **11**, 42–65 (2009).
2. EGAPP Working Group. Recommendations from the EGAPP Working Group: genetic testing strategies in newly diagnosed individuals with colorectal cancer aimed at reducing morbidity and mortality from Lynch syndrome in relatives. *Genet. Med.* **11**, 35–41 (2009).
3. Rogowski, W. Genetic screening by DNA technology. A systematic review of health economic evidence. *Int. J. Technol. Assess Health Care* **22**, 327–337 (2006).
4. Grossman, I. Routine pharmacogenetic testing in clinical practice: dream or reality? *Pharmacogenomics* **8**, 1449–1459 (2007).
5. Krawczak, M., Caliebe, A., Croucher, P. J. & Schmidtke, J. On the testing load incurred by cascade genetic carrier screening for Mendelian disorders: a brief report. *Genet. Test.* **11**, 417–419 (2007).
6. Schmidtke, J., Pabst, B. & Nippert, I. DNA-based genetic testing is rising steeply in a national health care system with open access to services: a survey of genetic test use in Germany, 1996–2002. *Genet. Test.* **9**, 80–84 (2005).
7. Khoury, M. J. *et al.* The evidence dilemma in genomic medicine. *Health Aff. (Millwood)* **27**, 1600–1611 (2008).
8. Rogowski, W., Grosse, S. D., John, J. & Palmer, S. The value of cost-effectiveness information for the decision on genetic screening for haemochromatosis in Germany. *University of York, Centre of Health Economics [online]*, <<http://www.york.ac.uk/inst/che>> (in the press).
9. Murray, T. H. in *Genetic Secrets: Protecting Privacy and Confidentiality in the Genetic Era* (ed. Rothstein, M.) 60–73 (Yale Univ. Press, London, 1997).
10. Teutsch, S. M. *et al.* The Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Initiative: methods of the EGAPP Working Group. *Genet. Med.* **11**, 3–14 (2009).
11. Ioannidis, J. P., Ntzani, E. E., Trikalinos, T. A. & Contopoulos-Ioannidis, D. G. Replication validity of genetic association studies. *Nature Genet.* **29**, 306–309 (2001).
12. Grosse, S. D. & Khoury, M. J. What is the clinical utility of genetic testing? *Genet. Med.* **8**, 448–450 (2006).
13. Payne, K. *et al.* Outcome measurement in clinical genetics services: a systematic review of validated measures. *Value Health* **11**, 497–508 (2008).
14. Foster, C. *et al.* Predictive genetic testing for BRCA1/2 in a UK clinical cohort: three-year follow-up. *Br. J. Cancer* **96**, 718–724 (2007).
15. Beery, T. A. & Williams, J. K. Risk reduction and health promotion behaviors following genetic testing for adult-onset disorders. *Genet. Test.* **11**, 111–123 (2007).
16. Rogowski, W. H. The cost-effectiveness of screening for hereditary hemochromatosis in Germany: a remodeling study. *Med. Decis. Making* **29**, 224–238 (2009).
17. Beran, T. M. *et al.* The trajectory of psychological impact in BRCA1/2 genetic testing: does time heal? *Ann. Behav. Med.* **36**, 107–116 (2008).
18. Smith, A. W. *et al.* Psychological distress and quality of life associated with genetic testing for breast cancer risk. *Psychooncology* **17**, 767–773 (2008).
19. Burnett, L., McQueen, M. J., Jonsson, J. J. & Torricelli, F. IFCC position paper: report of the IFCC taskforce on ethics: introduction and framework. *Clin. Chem. Lab. Med.* **45**, 1098–1104 (2007).
20. Sugarman, J. & Sulmasy, D. P. *Methods in Medical Ethics* (Georgetown Univ. Press, Washington, 2001).
21. Potter, B. K. *et al.* Guidance for considering ethical, legal, and social issues in health technology assessment: application to genetic screening. *Int. J. Technol. Assess. Health Care* **24**, 412–422 (2008).
22. Javaher, P. *et al.* EuroGentest: DNA-based testing for heritable disorders in Europe. *Community Genet.* **11**, 75–120 (2008).
23. Rogowski, W. Current impact of gene technology on healthcare. A map of economic assessments. *Health Policy* **80**, 340–357 (2007).
24. McGuire, A. L. & Burke, W. An unwelcome side effect of direct-to-consumer personal genome testing: raiding the medical commons. *JAMA* **300**, 2669–2671 (2008).
25. Chikhaoui, Y., Gelinas, H., Joseph, L. & Lance, J. M. Cost-minimization analysis of genetic testing versus clinical screening of at-risk relatives for familial adenomatous polyposis. *Int. J. Technol. Assess. Health Care* **18**, 67–80 (2002).
26. Joseph, B., Shanmugam, M. P., Srinivasan, M. K. & Kumaramanickavel, G. Retinoblastoma: genetic testing versus conventional clinical screening in India. *Mol. Diagn.* **8**, 237–243 (2004).
27. Grosse, S. D., Wordsworth, S. & Payne, K. Economic methods for valuing the outcomes of genetic testing: beyond cost-effectiveness analysis. *Genet. Med.* **10**, 648–654 (2008).
28. Dodge, J. A., Lewis, P. A., Stanton, M. & Wilsher, J. Cystic fibrosis mortality and survival in the UK: 1947–2003. *Eur. Respir. J.* **29**, 522–526 (2007).
29. Claxton, K., Sculpher, M. & Drummond, M. A rational framework for decision making by the National Institute For Clinical Excellence (NICE). *Lancet* **360**, 711–715 (2002).
30. Rogowski, W. H., Hartz, S. C. & John, J. H. Clearing up the hazy road from bench to bedside: a framework for integrating the fourth hurdle into translational medicine. *BMC Health Serv. Res.* **8**, 1–12 (2008).
31. Garber, A. M. Cost-effectiveness and evidence evaluation as criteria for coverage policy. *Health Aff. (Millwood)* [online], <<http://content.healthaffairs.org/cgi/content/short/hlthaff.w4.284>> (2004).
32. Taylor, R. S., Drummond, M. F., Salkeld, G. & Sullivan, S. D. Inclusion of cost effectiveness in licensing requirements of new drugs: the fourth hurdle. *BMJ* **329**, 972–975 (2004).
33. Vegter, S. *et al.* Pharmacoeconomic evaluations of pharmacogenetic and genomic screening programmes: a systematic review on content and adherence to guidelines. *Pharmacoeconomics* **26**, 569–587 (2008).
34. Boulenger, S. *et al.* Can economic evaluations be made more transferable? *Eur. J. Health Econ.* **6**, 334–346 (2005).
35. Carlson, J. J., Henrikson, N. B., Veenstra, D. L. & Ramsey, S. D. Economic analyses of human genetics services: a systematic review. *Genet. Med.* **7**, 519–523 (2005).
36. Olsen, J. A. Theories of justice and their implications for priority setting in health care. *J. Health Econ.* **16**, 625–639 (1997).
37. Bleichrodt, H., Crainich, D. & Eeckhoudt, L. Aversion to health inequalities and priority setting in health care. *J. Health Econ.* **27**, 1594–1604 (2008).
38. Callahan, D. & Milbank Memorial Fund. *What Price Better Health? Hazards of the Research Imperative* (Univ. California Press, Berkeley, 2003).
39. Claxton, K. P. & Sculpher, M. J. Using value of information analysis to prioritise health research: some lessons from recent UK experience. *Pharmacoeconomics* **24**, 1055–1068 (2006).
40. Grosse, S. D. Assessing cost-effectiveness in healthcare: history of the \$50,000 per QALY threshold. *Expert Rev. Pharmacoecon. Outcomes Res.* **8**, 165–178 (2008).
41. Moyer, V. A., Calonge, N., Teutsch, S. M. & Botkin, J. R. Expanding newborn screening: process, policy, and priorities. *Hastings Cent. Rep.* **38**, 32–39 (2008).
42. Green, N. S. *et al.* Committee Report: advancing the current recommended panel of conditions for newborn screening. *Genet. Med.* **9**, 792–796 (2007).
43. Baily, M. A. & Murray, T. H. Ethics, evidence, and cost in newborn screening. *Hastings Cent. Rep.* **38**, 23–31 (2008).

44. Hurst, S. A. *et al.* Prevalence and determinants of physician bedside rationing: data from Europe. *J. Gen. Intern. Med.* **21**, 1138–1143 (2006).
45. Gold, M. R., Siegel, J. E., Russell, L. B. & Weinstein, M. C. (eds). *Cost-Effectiveness in Health and Medicine* (Oxford Univ. Press, New York, 1996).
46. Wasson, K., Cook, E. D. & Helzlsouer, K. Direct-to-consumer online genetic testing and the four principles: an analysis of the ethical issues. *Ethics Med.* **22**, 83–91 (2006).
47. Daniels, N. *Just Health: Meeting Health Needs Fairly* (Cambridge Univ. Press, Cambridge, 2008).
48. Phillips, Z., Bojke, L., Sculpher, M., Claxton, K. & Golder, S. Good practice guidelines for decision-analytic modelling in health technology assessment: a review and consolidation of quality assessment. *Pharmacoeconomics* **24**, 355–371 (2006).
49. Cassiman, J. J. Research network: EuroGentest — a European Network of Excellence aimed at harmonizing genetic testing services. *Eur. J. Hum. Genet.* **13**, 1103–1105 (2005).
50. Health Council of the Netherlands. Screening: between hope and hype. Publication no. 2008/05E. *The Hague: Health Council of the Netherlands* [online]. <<http://www.gr.nl/pdf.php?ID=1702&p=1>> (2008).
51. Walton, S. M. *et al.* Prioritizing future research on off-label prescribing: results of a quantitative evaluation. *Pharmacotherapy* **28**, 1443–1452 (2008).
52. Noorani, H. Z., Huseraue, D. R., Boudreau, R. & Skidmore, B. Priority setting for health technology assessments: a systematic review of current practical approaches. *Int. J. Technol. Assess. Health Care* **23**, 310–315 (2007).
53. Hutton, J., Trueman, P. & Henshall, C. Coverage with evidence development: an examination of conceptual and policy issues. *Int. J. Technol. Assess. Health Care* **23**, 425–432 (2007).
54. Commission of the European Communities. Implementation of the Council Recommendation of 2 December 2003 on cancer screening (2003/878/EC). *European Commission* [online]. <[http://ec.europa.eu/health/ph\\_determinants/genetics/documents/com\\_2008\\_882\\_en.pdf](http://ec.europa.eu/health/ph_determinants/genetics/documents/com_2008_882_en.pdf)> (2008).
55. Ramsey, S. D., Burke, W. & Clarke, L. An economic viewpoint on alternative strategies for identifying persons with hereditary nonpolyposis colorectal cancer. *Genet. Med.* **5**, 353–363 (2003).
56. Grosse, S. D. *et al.* Population screening for genetic disorders in the 21st century: evidence, economics, and ethics. *Public Health Genomics* (in press).
57. Welte, R., Feenstra, T., Jager, H. & Leidl, R. A decision chart for assessing and improving the transferability of economic evaluation results between countries. *Pharmacoeconomics* **22**, 857–876 (2004).
58. Phillips, K. A. Closing the evidence gap in the use of emerging testing technologies in clinical practice. *JAMA* **300**, 2542–2544 (2008).
59. Wolff, A. C. *et al.* American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. *J. Clin. Oncol.* **25**, 118–145 (2007).

#### Acknowledgements

We acknowledge the comments from various colleagues, especially M.-A. Bailey, D. Callahan, L. Friedman Ross, J. John, R. Leidl and J. Schmidtke. All errors of omission and commission remain our own. W.R. was partly supported by EuroGentest, an EU-FP6 supported NoE contract number 512148 (EuroGentest Unit 3: Clinical genetics, community genetics and public health, Workpackage 3.2 (J. Schmidtke)). The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the US Centers for Disease Control and Prevention.

#### FURTHER INFORMATION

Wolf Rogowski's homepage:

<http://www.helmholtz-muenchen.de/en/igm/institute/team/staff-members/wolfrrogowski/index.html>

Muin J. Khoury's homepage:

<http://www.cdc.gov/genomics/about/mkhoury.htm>  
 ACCE: <http://www.cdc.gov/genomics/qTesting/ACCE.htm>

EuroGentest: <http://www.eurogentest.org>

Evaluation of Genomic Applications in Practice and

Prevention (EGAPP): <http://www.egappreviews.org>

National Institute of Health and Clinical Excellence

(NICE): <http://www.nice.org.uk>

ALL LINKS ARE ACTIVE IN THE ONLINE PDF

#### OPINION

## Mitochondrial and plastid evolution in eukaryotes: an outsiders' perspective

Jeferson Gross and Debashish Bhattacharya



Darwin200

Abstract | The eukaryotic organelles mitochondrion and plastid originated from eubacterial endosymbionts. Here we propose that, in both cases, prokaryote-to-organelle conversion was driven by the internalization of host-encoded factors progressing from the outer membrane of the endosymbionts towards the intermembrane space, inner membrane and finally the organelle interior. This was made possible by an outside-to-inside establishment in the endosymbionts of host-controlled protein-sorting components, which enabled the gradual integration of organelle functions into the nuclear genome. Such a convergent trajectory for mitochondrion and plastid establishment suggests a novel paradigm for organelle evolution that affects theories of eukaryogenesis.

A hallmark of eukaryotic cells is the presence of a nucleus and the bioenergetic organelles — the mitochondrion and the plastid, which originated from Gram-negative bacterial progenitors that once resided inside a eukaryotic 'host' cell as endosymbionts and were converted over time into bona fide organelles<sup>1–6</sup>. It is believed that the mitochondrion originated through endosymbiosis from an alphaproteobacterium approximately two billion years ago<sup>1–6</sup>, but the nature of the host remains uncertain and lies at the heart of an ongoing debate about eukaryogenesis<sup>3</sup>. One line of reasoning argues that an endomembrane system was a prerequisite for phagocytic capture of the alphaproteobacterium, implying that the host of the mitochondrial forerunner was a primitive eukaryote<sup>3,6</sup>. The opposing view is that the eukaryotic stem lineage was founded by the association of the alphaproteobacterial endosymbiont with a methanogenic archaeon, with the endosymbiont exerting a prominent influence on the emergence of the endomembrane system and the nuclear envelope in the methanogenic host<sup>7,8</sup>.

Despite these divergent opinions about mitochondrial origin, there is broad agreement that the plastid originated approximately 500 million years later in a bona fide mitochondriate eukaryotic host through a cyanobacterial primary endosymbiosis<sup>1,2,9</sup>. Although the phylogenetic provenance of the cells (if not the nature of the host) that

gave rise to the mitochondrion and the plastid has been clarified, a satisfactory explanation of how the prokaryote-to-organelle conversion (organellogenesis) occurred is lacking. However, it is clear that this process led to a strict dependence of the organelles on the nucleus<sup>1,2,4–6,10–12</sup>. This is substantiated by the fact that over 2,000 proteins that function in modern day mitochondria and plastids are encoded in the nuclear genome, synthesized by cytosolic ribosomes and then transported into the endosymbiont-derived compartments<sup>13,14</sup>. Such a nucleus-to-organelle flow of information constitutes the essential property of organellogenesis that resulted from the eukaryotic host progressively assuming control over the biogenesis of the captured prokaryotic cells.

Nuclear dominance over bioenergetic organelles is probably the result of continuous selective pressure on the host to optimize energy production from alphaproteobacterial oxidative phosphorylation and cyanobacterial oxygenic photosynthesis<sup>1,2,11,12</sup>. Therefore, innovations that increased the efficiency of energy production and its extraction from the endosymbionts were fixed over time in the nuclear genome. In this context, it is likely that the genomes of the mitochondrial and plastid forerunners were degenerating owing to Muller's ratchet<sup>10,11,15</sup>, resulting in a decrease in bioenergetic performance. To maintain and further improve the physiological competence of the endosymbionts, the