public health genomics
European Network

A DG SANCO issued European Agency for Health and Consumers Project

www.phgen.eu

European Best Practice Guidelines for Quality Assurance, Provision and Use of Genome-based Information and Technologies

PART I
European Best Practice Guidelines for Quality Assurance of Genome-based Information and Technologies

April 2012

EU Project No. 20081302 June 2009 – November 2012

PHGEN Interim Report of the Guidance for Quality Assurance of Genome-based Information and Technologies

Karla Douw1, Hindrik Vondeling1, Jean-Jacques Cassiman2, Elfriede Swinnen2, Pascal Borry3, and Heidi Howard3

1 Department of Health Technology & Services Research (HTSR), University of Twente, The Netherlands
2 Department of Human Genetics, Katholic University Leuven, Belgium
3 Interfaculty Centre for Biomedial Ethics and Law, Katholic University Leuven, Belgium.

Contents

Foreword

The structure of this report in recommendations and best practice according to the public health wheel was provided by the coordinating group of the project. Also the operationalization of the wheel was done by this group. Quality assurance of genome-based technologies touches upon many of the tasks in the wheel, but not on all. The working group's definition of quality assurance can be found in paragraph 2.1. It was defined at the start of the project and based on the objective in the original project description.

1. Executive summary

The application of genome-based technologies and information with the aim of combating diseases of public health significance still holds a great promise, but these advances also bring with them a slew of ethical and social issues that challenge the normative frameworks used in clinical genetics until now. This report outlines five principles that should be considered in order to responsibly introduce genome based technologies and information into public health. On a more operationalized level, the report outlines the various analytic frameworks that are available to assess and appraise genome-based technologies and information, and provides recommendations for further development and implementation.

2. Specification of the pillar / working definition

A best practice guideline for Quality Assurance in the field of public health genomics is aimed at supporting the translation and application of genome-based technology and information for the purpose of combating diseases of public health significance. Chou et al. (1) state that until now "quality measurement in genetics has focused more on patient satisfaction and the presence or absence of program components… rather than effectiveness of care and its impact on health outcomes." Ibarreta et al. (2) conclude for the European context, that establishment of the clinical utility of the new type of tests is one of the gaps in current quality assurance schemes. The focus on health outcomes is becoming more important, as the number of new genetic tests and genomic tests for common complex diseases is increasing, and their effectiveness in improving health and health services is not (yet) clear.

Recently published guidelines and reviews underpin the necessity of guiding the assessment of genome-based discoveries and of supporting decisions on how to use the results in clinical practice (3).

Against this background, Quality Assurance of Genome-based technology and information (GBT and GBI) will be defined in relation to their responsible introduction into health care.

2.1 General objective of the pillar: (perspective / definition)

Quality assurance in this context is defined as guidance for the correct and responsible application of genome-based technology (GBT) and genome-based information (GBI) in public health in terms of their validity, utility and efficiency.

NB. QA in the context of PHGEN II does not include the evaluation of the capabilities of public health staff working with these technologies, nor the settings (e.g. laboratories, hospitals, practices) in which the technologies are developed and used.

2.2. Specific objectives of the pillar (target audience)

The pillar has the following objectives:

a. To identify existing analytical frameworks for the introduction of genome based information and technologies

b. To identify normative frameworks for the introduction of genome-based information and technologies.

c. Identify gaps and formulate recommendations for new guidance

A focused study on the ethical, legal and social issues (ELSI) that arise from the use of GBT and GBI in health care and public health will be performed to propose a normative framework for the introduction of GBT and GBIs. For this purpose the genetic/genomic ELSI literature will be gathered and reviewed in order to identify the relevant issues

The framework will provide the overarching criteria for a discussion of the existing guidance on identifying, selecting and assessing GBT and GBIs, by means of analytic frameworks. Rosenkoetter et al. (3) state in their paper that (…) HTA (in combination with CTA) offers the most comprehensive set of methods for assessment of Public Health Genomics'. Starting point for our work is to compare existing analytic

frameworks (used in the context of HTA), and to analyse them on the basis of content analysis.

A literature review will be performed to add additional notions on evaluation of GBTs to the existing frameworks.

Inclusion of notions of abovementioned analyses will be translated into meta-guidance, i.e. referral to existing guidance with addition of recommendations for new guidance for identification, selection and assessment of GBTs for public health.

Target audience

The target audience of the guidance will be policy makers at the macro level of (public) health care systems.

2.3. Working group members and task allocation

The working group consisted of a core group and group members. The core group was responsable for meeting the working groups objectives, and the group members (had) a supporting role in terms of providing information on relevant sources, stakeholders, and in reviewing documents produced by the core group. The core group existed of:

Hindrik Vondeling, Department of Health Technology & Services Research (HTSR), University of Twente, Netherlands

Karla Douw, Department of Health Technology & Services Research (HTSR),

University of Twente, Netherlands

Jean-Jacques Cassiman, Department of Human Genetics, Katholic University Leuven, Belgium

Elfriede Swinnen, Department of Human Genetics, Katholic University Leuven, Belgium

The group worked together with:

Pascal Borry, and Heidi Howard, Interfaculty Centre for Biomedial Ethics and Law, Katholic University Leuven, Belgium.

The Working group had the following group members:

Iñaki Gutiérrez Ibarluzea, Department of Health, Osteba, Basque Office for HTA, Spain

Stefania Boccia, Institute of Hygiene and Preventive Medicine, Catholic University of the Sacred Rome, Italy

Hans Westerhoff, Vrije Universiteit Amsterdam, The Netherlands

Miroslaw Wysocki, National Institute of Public Health, Poland

Núria Malats, Fundación Centro Nacional de Investigaciones Oncológicas Carlos III, Spain.

2.4. General background documentation of QA pillar

The focus of the QA pillar is on the introduction of genome-based technology and information in public health. This includes issues of safety, accountability, and equity in adoption and use of genome-based technology and information in health care. The constantly rising interest in, and the push for genetic and genomic testing for common complex disorders as well as the actual offer of genetic and genomic testing for such disorders (also by commercial companies) are seriously challenging the analytical and normative frameworks previously used in clinical genetics. Furthermore, the shift from the public health focus on populations to that of the individual, also formerly known as patient or consumer, makes it necessary to

discuss and analyse how adequate the existing frameworks are, and to what extent they have to be adapted in order for GBT and GBI to be appropriately translated into health care.

2.4.1 Normative framework for introduction of GBT and GBIs

The great importance of addressing the ethical, legal and social issues of genetic and genomic science is widely accepted and promoted by a wide range of stakeholders. Since the establishment of the ELSI Working Group for the human genome program in 1989, a large corpus addressing ethical, legal and social issues arising from genetic and genomic technologies and services has been produced by experts from many different fields (genetics, medicine, law, philosophy etc.). This work has helped guide, among others, many aspects of genetic test and information delivery for patients as well as issues in genetic and genomic research. The larger themes for which ELSI will be discussed have been identified from the chapter entitled “Criteria for responsible screening” developed by the Health Council of the Netherlands in their report “Screening: between hope and hype” (4). Although it is realized that the scope of use of GBTs goes beyond one single area such as (genetic/genomic) screening, this outline of themes will be used as an initial framework from which to build up the ELSI discussion. Five themes are distinguished.

(1) Reliable and valid process and reliable and valid genome based interventions

The introduction of GBI into health care and public health should be based on a solid scientific foundation. Furthermore, a framework and process for this translation should be established before the actual introduction of the GBT and GBIs and the quality of the various parts of the process must be monitored throughout the period in which the GBI are offered to the public.

An important issue for this principle, as well as a recurring theme throughout this document is the answer to the following questions: 1) *What* will be the criteria (and

the relevant thresholds) for reliability and validity? And 2) Who will be responsible to decide these criteria and implement the processes involved?

2) GBI introduced into the health care system and financed by public means should be focused on significant health problems.

In the context of limited financial resources for health services, applying to all EU member states , the notion of prioritizing is paramount. For these reasons, we agree that GBI should focus on important or significant health problems. *What is important?*: As mentioned above, complicating matters here are what *are* the specific criteria for determining a "very serious condition" and *who* should decide? In answering the question of what is important, one should keep in mind that it is often the case with new technologies that they are introduced not on the basis of a health or health care service need, but on the basis of a technological imperative.

3) The advantages of introducing and offering GBTI should outweigh the disadvantages.

The principle questions here are what are the criteria needed to be included in the calculation of clinical utility in the broad sense (or benefits and risks) and *who* will take the responsibility of performing the calculation. Both variables have an important impact on the nature and quality of the answer. Transparency regarding the actual measures used in the calculation, as well as the underlying value judgements and contextual factors are important.

4) The autonomy of patients and individuals in general, must be respected

The different approach in genomics, compared to the traditional findings in genetics, challenges the continuum between autonomy, public welfare and Public Health. While genetic risks in diseases like Huntington are unavoidable, gene-environment interactions encompass different risk factors. Some of the risk factors are modifiable. E.g the development of "genomic" diseases like Diabetes type II is influenced by many factors including the diet, alcohol consumption and physical activity. Some of the scenarios which have dominated the discourse in the past, like the conflict between "the right to know" and the "right not to know", will be of minor importance as the concept of determinism fades away in many disease clusters (5). Furthermore, the discussion of respect for autonomy of a patient in the context of a clinical encounter between a patient and a doctor differs greatly from that in the context of a public health program such as genetic screening or vaccination. For public health interventions, the actions may involve many (or mostly) healthy people and "require something approaching certainty as to the benefits and possible side effects of an intervention". Regardless of the context, the test must be acceptable to the patient and/or to the target population and participation should be voluntary.

The content and way in which the necessary information on GBT and GBIs is provided in both scenarios can vary greatly. Since we are discussing herein the introduction of GBI into the health care system in general (i.e. including both genomic testing and genomic screening) it is primordial that both versions of autonomy be anticipated and a framework be put into place for it to be respected appropriately in both circumstances.

5) The offer of GBI funded from public sources should be justified in the context of the overall healthcare budget.

There are three main aspects in this criteria. Firstly, GBTI programmes, just like any other health intervention, need to be affordable in terms of overall spending relative to the size of the budget, be it the total health care budget or, the more limited

budget for screening. This differs per country and is hardly ever explicitly stated. Secondly, GBTI programmes should offer value for money, that is, the extra health gain of a programme should justify the additional health care costs involved. Another important aspect, not mentioned in the Health Council document, is the role of equity considerations and the trade-offs that can be made between equity and efficiency in allocation of scarce resources. Equity can then either be defined as equity in access to care, equity in health care consumption, or equity in health outcomes. As yet, there is no well-accepted methodology to systematically include these considerations in decision making.

Public versus private sector

An important question remains: should the five criteria discussed above also be applied to the private health sector wherein services are not funded by the state? This is a relevant question since presently genome testing is being offered directly to the public by private commercial companies (such as 23andMe), often completely bypassing the health care system and the intervention of a health care professional. Should these GBI also be subject to the criteria for responsible introduction of tests? We believe that of the five criteria described above, at least the following three should be applied in the private health care sector: ensuring a reliable and valid process of introduction and application of GBI; the benefits should outweigh the disadvantages; and there should be respect for autonomy. The other two criteria involving the focus on an important/significant health problem and a responsible use of public resources are less relevant to the private sector, yet should still be considered in certain contexts.

The abovementioned criteria constitute the *normative framework* for a discussion on how to identify, select and assess GBTs, by means of *analytic* frameworks.

2.4.2 Analytic framework for introduction of GBT and GBIs

Background on introduction of GBT and GBIs

In spite of accelerating human genome discoveries in a wide variety of diseases of public health significance, the promise of personalized health care and disease prevention based on genomics has lagged behind (6). Khoury introduced a continuum to illustrate the translation in different phases from discovery to application (see figure 1).

T1 T2 T3 T4

Gene discovery < > Health Application < > Health Practice < > Health Impact

Figure 1 The translational process (based on Khoury et al. (7))

There are a number of factors influencing the flow in this continuum, and for the pillar of QA the factors influencing the phases T2 to T4 are important. The processes that influence the pace of the translation process in these phases are amongst other, regulatory policies, clinical practice liability issues, oversight, clinical practice guidelines, coverage and reimbursement (8). Rogowski et al. (9) have illustrated how the translational process looks like with the addition of the factor 'coverage and reimbursement' (see figure 2).

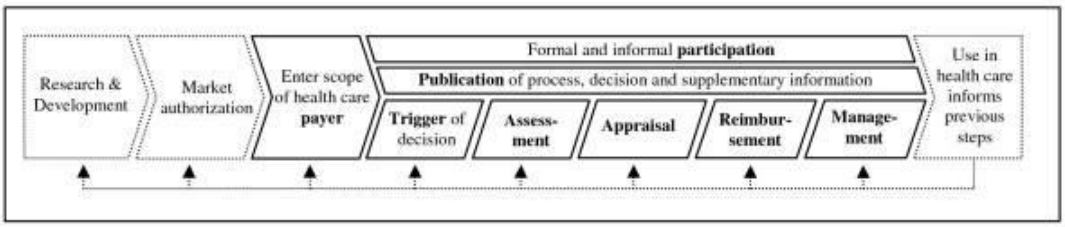


Figure 2. The fourth hurdle within the process of translational medicine (source: 9)

In all of these processes the evidence available on the potential impact of the GBTI plays a role, as the evidence needs to be summarized to answer specific questions either in the context of national regulatory policies, or for local context of e.g. the development of a clinical guideline. Analytic frameworks are used for this purpose, to transparently organize the summarizing of evidence in relation to specific questions regarding the introduction of the technology (10).

Currently, where there is very little translational research going on on these discoveries, and thus very little data available on the value of the GBT and GBIs, two situations can occur. One is that the regulator or adopter applies a *low threshold for translation into practice*. This results in a short T2 phase according to Khoury (8), and early introduction into practice. In this situation GBT and GBIs are introduced in practice while there is little information on clinical validity, no information on clinical utility, and therefore uncertainty about the potential for increased harms and for increased benefits, and use will be based on expert opinion.

The other situation is one in which the regulator or adopter applies a *high threshold for translation into practice*. This results in a long T2, and a late introduction. In this situation, valid and useful tests are introduced in practice, on the basis of a strong evidence base. There will be a diminished potential for harms, but there is potential for diminished benefits, in that the technology was introduced later than it could have been, so certain patients or consumers might have missed out on its benefit.

Responsible introduction of new health technologies into the health care system will always have to deal with this trade-off between a rigorous evaluation based on sufficient evidence on the new health technology and the timing of introduction into health care. In the field of Health Technology Assessment this dilemma has also been coined: "It's always too early to evaluate a technology… …until suddenly it's too late." by *Professor Martin Buxton, Brunel University, London, UK [Buxton's Law].*

It is important to make clear that this dilemma will always exist, and that it is up to the regulator or adopter of the technology to make the trade-off. Ideally, the choice

is transparent and accountable. The choice can be made transparent by systematically following the steps in the process of HTA.

Health Technology Assessment

HTA is a multidisciplinary process that summarises information about the medical, social, economic and ethical issues related to the use of a health technology in a systematic, transparent, unbiased, robust manner. Its aim is to inform the formulation of safe, effective, health policies that are patient focused and seek to achieve best value (11).

The process of HTA exemplifies the trade-off between rigor and timeliness for regulators and adopters in that it comprises of 6 interdependent types of actions (12), which are aimed at finding those new health technologies that are of most significance for the adopter or regulator:

1) identification of new health technologies (Horizon Scanning);

2) priority setting, selecting those technologies most in need of assessment;

3) testing, conducting the appropriate data collection, and data analysis;

4) synthesis, collecting and interpreting existing information and the results of the testing step and, usually, making recommendations or judgements about appropriate use;

5) dissemination, providing the synthesized information, or any other relevant information, to the appropriate persons who use or make decisions concerning the use of health technologies;

6) implementation, assuring that changes in knowledge and attitudes result in changes in behaviour or in particular decisions.

2.4.3 Existing guidance and best practice

Identification of Genome-based technologies and information

There is general guidance available on how to set up a systematic approach for Horizon Scanning of potentially significant new health technologies. This is provided by EuroScan. Furthermore, there is experience at the U.S. CDC's Office of Public Health Genomics with scanning for genome-based technology and information (see below).

EuroScan is a collaborative network of member agencies for the exchange of information on important emerging new drugs, devices, procedures, programs, and settings in health care. The members produced a *Toolkit for the identification and assessment of new and emerging health technologies* (13) in 2009. It can be obtained from EuroScan's web site http://euroscan.org.uk/methods/

The toolkit can be considered as a consensus document on what constitutes Horizon Scanning of new health technologies, i.e. the steps involved, and the related best practices for identification, selection and assessment of new health technologies. It provides guidance on the various stages and provides questions that anyone establishing or improving an Early Awareness and Alert systems (EAAS) should ask themselves), i.e. on identifying the market; determining the time horizon; identification, filtration, prioritization and assessment of health technologies; dissemination of the information and updating the information. The toolkit does not represent best evidence on these methods, but represents best practice in the sense of generally-accepted, informally-standardized techniques, methods, and processes that are advocated by the existing international network for Horizon Scanning in the public domain.

Gwinn et al. (14) describes the development and testing of a search strategy for newly developed or launched genomic tests for the Evaluation of Genomic Applications in Practice and Prevention (EGAPP) project. The search is based on weekly Google Alert queries, and the results are entered into an on-line searchable database GAPP finder (http://www.hugenavigator.net/GAPPKB/topicStartPage.do). GAPP finder thus provides a source in itself for others involved in identifying new genome-based technologies. The tests are categorized in the database according to *the disease or condition* (e.g. irritable bowel syndrome), *the type of test* (e.g. expression of 32 genes in tumor tissue), *the target population* (e.g women with a family history of ovarian cancer), *and the intended use* (e.g. appropriate warfarin dosing).

Priority Setting for assessment of GBT and GBIs

Criterion 2 of the normative framework stated that '*GBT and GBIs introduced into the health care system and financed by public means should be focused on significant health problems.'* In priority setting for HTA this issue of 'significant health problem' is generally reflected in a criterion on the burden of disease (15). For rare diseases often different priority setting procedures exist, and other criteria are used, such as rule of rescue. Andermann et al. (16) have added that in relation to genetic screening one can distinguish a population viewpoint and a patient viewpoint. From the first perspective the focus will be on common health problems, and priorities have to be defined on an operationalization of burden of disease in that case. And from the patient viewpoint one focuses on serious health problems, which have to be defined in terms of morbidity, mortality and disability. Exemplifications of Not serious health problems include e.g. non-medical conditions, and sex selection. This distinction between a population and patient viewpoint will be important in relation to discussion about allocation of resources for assessment and use of GBT and GBIs.

Priority setting for assessment is not the same as priority setting for reimbursement or coverage. The subgroup on priority setting of the EURASSESS project (17) stated that the aim of setting priorities for HTA should be to identify those assessments that offer the greatest benefit in relation to their cost, and thus to maximise the benefit derived from investment in HTA. The aim is to identify the benefits that will result from an assessment and not (necessarily) the benefit that may result from the technology in question. The preferred focus and leading principle for priority setting is that an assessment should lead to a reduction in uncertainty, and as such serve as an aid to decision makers. Henshall et al. (17) explain priority setting for assessment as follows 'Assume, for example, that a technology is believed to offer substantial benefits at modest cost to a large number of people concerned and is therefore widely adopted. The benefit of assessing the technology will depend upon the degree of uncertainty about the actual effects of the technology. If there is little doubt about its effects, there will be little value in assessing it, even though the technology is of substantial value to a large number of people. There may be more benefit in assessing a technology of lesser value but associated with greater uncertainty'.

This principle should also be applied to GBT and GBIs assessment with public resources. Those technologies where there is greatest uncertainty of their benefit, should be selected.

On the whole, the same set of priority setting criteria for HTA is used internationally, although there are slight variations, there are certainly differences in how criteria are weighted (15;18). A recent overview of possible criteria is provided by the EU-funded project, Health Investment for Screening and Diagnosis of disease project, which focused on genetic/genomic testing (NB final document not yet issued, work in progress):

- potential for improved patient benefits, including health gains but also, non health gains such as the value of information.

- availability of other (similar) technologies;
- cost and potential budget impact of the technology;
- extent of evidence available for making a decision, or, making recommendations for future research required;
- level of uncertainty in the data;
- risk of harm from not conducting a HTA;
- potential for the HTA to make an impact, such as meeting 'unmet need' and adding value to the healthcare system and patients;
- and specifically for pharmacogenetics, the cost of the drug to be prescribed or the severity of the potential side effects being avoided by targeting prescription.

EGAPP has produced criteria for preliminary ranking of topics (19), which partly resemble the other priority setting criteria. The following criteria are used by EGAPP:

Criteria related to health burden

What is the potential public health impact based on the prevalence/incidence of the disorder, the prevalence of gene variants, or the number of individuals likely to be tested?

What is the severity of the disease?

How strong is the reported relationship between a test result and a disease/drug response?

Is there an effective intervention for those with a positive test or their family members?

Who will use the information in clinical practice (e.g., healthcare providers, payers) and how relevant might this review be to their decision-making?

Criteria related to practice issues

What is the availability of the test in clinical practice?

Is an inappropriate test use possible or likely?

What is the potential impact of an evidence review or recommendations on clinical practice? on consumers?

Other considerations

How does the test add to the portfolio of EGAPP evidence based reviews? As a pilot project, EGAPP aims to develop a portfolio of evidence reviews that adequately tests the process and methodologies.

Will it be possible to make a recommendation, given the body of data available? EGAPP is attempting to balance selection of somewhat established tests versus emerging tests for which insufficient evidence or unpublished data are more likely.

Are there other practical considerations? For example, avoiding duplication of evidence reviews already underway by other groups.

How does this test contribute to diversity in reviews? In what category is this test? As a pilot project, EGAPP aims to consider different categories of tests (e.g., pharmacogenomics or cancer), mutation types (e.g., inherited or somatic) or test types (e.g., predictive or diagnostic).

A principle that is used by e.g. the Dutch Health Care insurance board is that the use of GBT or GBI needs to be followed by a treatment, and that this test-treatment chain is the starting point for the assessment of value. The treatment can be defined

broadly in terms of any intervention that follows after the test that improves patient outcomes. This would lead to the situation that the use of a new GBT that establishes the probability for a serious health problem, for which no treatment is available should not be prioritized with public means.

Improved patient outcome relates to the patient benefit criterion, and includes taking into account the clinical utility of GBT and GBIs. In addition to clinical utility the notion of personal utility (such as advantages of learning about genomic risk) was introduced (20;21). This notion might be taken into account in priority setting, i.e. included in the patient benefit criterion, but this is subject of discussion. Foster et al. (22) state that 'For purposes of using limited third party or public health resources, utility should be evaluated in a societal rather than individual context. However, for other health uses of genomic information a broader sense of overall utility should be used. Behavioral changes and increased individual awareness of health-related choices are relevant metrics for evaluating the personal utility of genomic information, even when traditional clinical benefits are not manifested.'

As is the case for all other criteria in a transparent process, the weighting of personal utility in relation to other criteria should also be made explicit.

Testing

In the testing phase evidence is gathered around the specific application of GBT and GBIs, and this can be organised by means of analytic frameworks. Analytic frameworks a) describe the context in which the technology is to be used, and b) to determine the nature of the evidence necessary for addressing questions around the technology (e.g. safety, effectiveness, cost-effectiveness, etc.) (10). The analytic frameworks structure the analytic process and identify gaps and weaknesses in the evidence.

Several analytic frameworks are available for evidence generation and synthesis of genetic and medical testing, and for health technologies in general. These include e.g. the ACCE framework, the Fryback-Thornbury framework, The US Preventive Services Task force framework for screening topics and counseling interventions (10), framework for evaluating evidence on tests (23), and more general the HTA core model for diagnostic and screening technologies (24).

A comparison of the ACCE framework and the HTA core model showed that the ACCE model lacks comparing the new technology to an alternative (existing) technology. Bossuyt (25) states that analytic framework for evaluating medical tests and markers will be more useful if they are comparative, 'i.e. explore the consequences of introducing of using the test versus the alternative, and focus on health outcomes.

Two central tenets emerged from the review of analytic frameworks.

(1) Evaluation of the value of a GBI must always be linked to the context of use; and (2) GBT and GBIs, and assessments of GBT and GBIs, are ultimately aimed improving the health outcomes of patients. This is represented in the causal chain diagram of figure 3.

GBI -> Result -> Categorization (e.g., high risk, disease present, disease progression) -> Decision -> Patient outcome

Figure 3. Causal chain diagram

GBT and GBIs will be used, just like medical tests in a complex context. The AHRQ (10) describes this context as including preexisting conditions, results of other tests, skill and knowledge of providers, availability of therapeutic resources, and so on. 'In this complex environment, researchers have tended to focus on narrow questions, such as the ability of a test to conform to technical specifications, to accurately classify patients into diagnostic or prognostic categories, or to influence thought or actions

by clinicians and patients. Rarely are medical tests evaluated in randomized controlled trials with representative patient populations and comprehensive measures of patient-relevant outcomes.' Arguably, the most robust empirical demonstration of the utility of a medical test is through a properly designed randomized controlled trial (RCT) that compares patient management strategies. As a result, the synthesis of the evidence resembles putting the pieces of a puzzle together (10).

This lack of comparative outcomes data is the same in the field of genome applications. Khoury et al. (26) blame regulatory and reimbursement policies that do not require such studies and also their inherent costs. And adds that there is no consensus on evidentiary requirements for genomic test evaluation, with some stakeholders accepting the findings of observational studies or even biological plausibility of potential benefits, whereas others insist on randomized controlled clinical trials (26).

The Evaluation of Genetic Tests and Genomic Applications group (EGAPP) has produced guidance to construct a "chain of evidence" , when direct evidence, i.e. evidence that directly relates the GBI to a health outcome is not available. After the overarching question, the remaining key questions address the components of evaluation as links in a possible chain of evidence: *analytic validity* (technical test performance), *clinical validity* (the strength of association that determines the test's ability to accurately and reliably identify or predict the disorder of interest), and *clinical utility* (balance of benefits and harms when the test is used to influence patient management). Determining whether a chain of indirect evidence can be applied to answer the overarching question of requires consideration of the quality of individual studies, the adequacy of evidence for each link in the evidence chain, and the certainty of benefit based on the quantity (i.e., number and size) and quality (i.e., internal validity) of studies, the consistency and generalizability of results, and understanding of other factors or contextual issues that might influence the conclusions (19).

Lijmers and Bosssuyt (27) conclude that the evaluation of medical tests is most likely to be a cyclic and repetitive process instead of a linear one, implying that studies of diagnostic accuracy do not always have to precede tests of clinical utility. Thus, the policy question related to the GBI defines which evidence ideally should be collected/available, and the chain of evidence therefore needs also to be related to the policy question.

In relation to the lack of evidence and uncertainty of evidence Veenstra et al. (28) suggest to apply a risk-benefit framework to assist the translation of GBT and GBIs. Most of the analytic frameworks mentioned above do not include the need for a reiterative process of assessment. Some, but not all include early assessment. Both aspects are however included in Early and constructiveTA, a form of technology assessment that starts early in the life cycle of development of a health technology, and in each stage defines the (cost)effectiveness of the technology relative to the context it will be used in. This also provides the possibility for adaptation during the R&D process of the technology, and the later stages. An example of a CTA in the area of genomics is that of genetic profiling for breast cancer by Retel et al. (29;30).

Risk-benefit assessment could be applied in an early TA, and would involve an iterative process, in that way it may accelerate the utilization and practice-based evidence development of genomic tests that pose low risk and offer plausible clinical benefit, while discouraging premature use of tests that provide little benefit or pose significant health risks compared with usual care (28).

As the evaluation of GBT and GBIs is aimed at supporting decision making by busy policy makers or local public health professionals, the application of a framework should lead to comprehensive, user-friendly, and timely reports on the available evidence, and recommendations for introduction. The two aspects of comprehensive and timely seem however not easily achievable in one framework. However, the need in practice is evident.

Synthesis, collecting and interpreting existing information and the results of the testing step and, usually, making recommendations or judgements about appropriate use

Currently, the best practice example on synthesizing the evidence of GBT and GBIs is provided by the EGAPP initiative. EGAPP has developed a hierarchy of quality of studies that studies analytic, and clinical validity and clinical utility (see table 1), which is mainly based on the methodology of the United States Preventive Services Task Force (USPSTF). These methods are currently tested by EGAPP itself, and will have to be tested by others as well, mainly against criteria of practicality, timeliness, and ease of use, as the information is aimed as supporting decision-making by busy health care professionals and policy makers.

Table 1 Hierarchies of data sources and study designs for the components of evaluation (taken from Teutsch et al. (19))

Level[a]	Analytic validity	Clinical validity	Clinical utility
1	Collaborative study using a large panel of well characterized samples	Well-designed longitudinal cohort studies	Meta-analysis of randomized controlled trials (RCT)
	Summary data from well-designed external proficiency testing schemes or interlaboratory comparison programs	Validated clinical decision rule[b]	
2	Other data from proficiency testing schemes	Well-designed case-control studies	A single randomized controlled trial
	Well-designed peer-reviewed studies (e.g., method comparisons, validation studies)		
	Expert panel reviewed FDA summaries		
3	Less well designed peer-reviewed studies	Lower quality case-control and cross- sectional studies	Controlled trial without randomization
		Unvalidated clinical decision rule[b]	Cohort or case-control study
4	Unpublished and/or non-peer reviewed research, clinical laboratory, or manufacturer data	Case series	Case series
	Studies on performance of the same basic methodology, but used to test for	Unpublished and/or non-peer reviewed research, clinical	Unpublished and/or non-peer reviewed

Level[a]	Analytic validity	Clinical validity	Clinical utility
	a different target	laboratory or manufacturer data	studies
		Consensus guidelines	Clinical laboratory or manufacturer data
		Expert opinion	Consensus guidelines
			Expert opinion

[a] Highest level is 1.
[b] A clinical decision rule is an algorithm leading to result categorization. It can also be defined as a clinical tool that quantifies the contributions made by different variables (e.g., test result, family history) in order to determine classification/interpretation of a test result (e.g., for diagnosis, prognosis, therapeutic response) in situations requiring complex decision-making.

Teutsch et al. (19) furthermore developed a checklist of questions for assessing the quality of individual studies for each evaluation component based on published literature (see table 2). 'The EWG ranks individual studies as *Good*, *Fair*, or *Marginal* based on critical appraisal using the criteria in Tables 1 and 2. The designation *Marginal* (rather than *Poor*) acknowledges that some studies may not have been "poor" in overall design or conduct, but may not have been designed to address the specific key question in the evidence review.' (19)

Table 2 Criteria for assessing quality of individual studies (internal validity) (taken from Teutsch et al. (19))

Analytic validity	Clinical validity	Clinical utility
Adequate descriptions of the index test (test under evaluation)	*Clear description of the disorder/phenotype and outcomes of interest*	*Clear description of the outcomes of interest*
Source and inclusion of positive and negative control materials	Status verified for all cases	What was the relative importance of outcomes measured; which were prespecified primary outcomes and which were secondary?
Reproducibility of test results	Appropriate verification of controls	*Clear presentation of the study design*
Quality control/assurance measures	Verification does not rely on *index test* result	Was there clear definition of the specific outcomes or decision options to be studied (clinical and other endpoints)?
Adequate descriptions of the test under evaluation	Prevalence estimates are provided	Was interpretation of outcomes/endpoints blinded?

Analytic validity	Clinical validity	Clinical utility
Specific methods/platforms evaluated	*Adequate description of study design and test/methodology*	Were negative results verified?
Number of positive samples and negative controls tested	*Adequate description of the study population*	*Was data collection prospective or retrospective?*
Adequate descriptions of the basis for the "right answer"	Inclusion/exclusion criteria	If an experimental study design was used, were subjects randomized? Were intervention and evaluation of outcomes blinded?
Comparison to a "gold standard" referent test	Sample size, demographics	Did the study include comparison with current practice/empirical treatment (value added)?
Consensus (e.g., external proficiency testing)	Study population defined and representative of the clinical population to be tested	*Intervention*
Characterized control materials (e.g., NIST, sequenced)	Allele/genotype frequencies or analyte distributions known in general and subpopulations	What interventions were used?
Avoidance of biases	*Independent blind comparison with appropriate, credible reference standard(s)*	What were the criteria for the use of the interventions?
Blinded testing and interpretation	Independent of the test	*Analysis of data*
Specimens represent routinely analyzed clinical specimens in all aspects (e.g., collection, transport, processing)	Used regardless of test results	Is the information provided sufficient to rate the quality of the studies?
Reporting of test failures and uninterpretable or indeterminate results	Description of handling of indeterminate results and outliers	Are the data relevant to each outcome identified?
Analysis of data	Blinded testing and interpretation of results	Is the analysis or modeling explicit and understandable?
Point estimates of analytic sensitivity and specificity with 95% confidence intervals	*Analysis of data*	Are analytic methods prespecified, adequately described, and appropriate for the study design?
Sample size/power calculations addressed	Possible biases are identified and potential impact discussed	Were losses to follow-up and resulting potential for bias accounted for?
	Point estimates of clinical sensitivity and specificity with 95% confidence intervals	Is there assessment of other sources of bias and confounding?
	Estimates of positive and negative predictive values	Are there point estimates of impact with 95% CI?
		Is the analysis adequate for the proposed use?

The next step is to grade the evidence for each step in the chain, i.e. analytic validity, clinical validity, and clinical utility, to define the quality of the total body of evidence

on the GBT. For this purpose, Teutsch et al. (19) developed a classification to assess the adequacy of the information to answer the key questions related to each evaluation component. The classification is either *Convincing*, *Adequate*, or *Inadequate* (see table 3). These questions are deemed critical to assess the "strength of linkages" in the chain of evidence (19).

Table 3 Grading the quality of evidence for the individual components of the chain of evidence (key questions) (taken from Teutsch et al. (19))

Adequacy of information to answer key questions	Analytic validity	Clinical validity	Clinical utility
Convincing	*Studies that provide confident estimates of analytic sensitivity and specificity using intended sample types from representative populations*	*Well-designed and conducted studies in representative population(s) that measure the strength of association between a genotype or biomarker and a specific and well-defined disease or phenotype*	*Well-designed and conducted studies in representative population(s) that assess specified health outcomes*
	Two or more Level 1 or 2 studies that are generalizable, have a sufficient number and distribution of challenges, and report consistent results	Systematic review/meta-analysis of Level 1 studies with homogeneity	Systematic review/meta-analysis of randomized controlled trials showing consistency in results
			At least one large randomized controlled trial (Level 2)
	One Level 1 or 2 study that is generalizable and has an appropriate number and distribution of challenges	Validated Clinical Decision Rule	
		High quality Level 1 cohort study	
Adequate	Two or more Level 1 or 2 studies that	Systematic review of lower quality studies	Systematic review with heterogeneity
	Lack the appropriate number and/or distribution of challenges	Review of Level 1 or 2 studies with heterogeneity	One or more controlled trials without randomization (Level 3)

Adequacy of information to answer key questions	Analytic validity	Clinical validity	Clinical utility
	Are consistent, but not generalizable	case/control study with good reference standards	Systematic review of Level 3 cohort studies with consistent results
	Modelling showing that lower quality (Level 3, 4) studies may be acceptable for a specific well- defined clinical scenario	Unvalidated Clinical Decision Rule (Level 2)	
Inadequate	Combinations of higher quality studies that show important unexplained inconsistencies	Single case-control study	Systematic review of Level 3 quality studies or studies with heterogeneity
		Nonconsecutive cases	
	One or more lower quality studies (Level 3 or 4)	Lacks consistently applied reference standards	Single Level 3 cohort or case-control study
	Expert opinion	Single Level 2 or 3 cohort/case-control study	Level 4 data
		Reference standard defined by the test or not used systematically	
		Study not blinded	
		Level 4 data	

This grading then needs to be translated into recommendations on the use of the technology in practice. EGAPP's main target group is clinicians, not policy makers.

Table 6 Recommendations based on certainty of evidence, magnitude of net benefit, and contextual issues (taken from Teutsch et al. (19))

Level of Certainty	Recommendation

Level of Certainty	Recommendation
High or moderate	Recommend for . . .
	. . . if the magnitude of net benefit is *Substantial, Moderate*, or *Small*[a], unless additional considerations warrant caution.
	Consider the importance of each relevant contextual factor and its magnitude or finding.
	Recommend against . . .
	. . . if the magnitude of net benefit is *Zero* or there are net harms.
	Consider the importance of each relevant contextual factor and its magnitude or finding.
Low	Insufficient evidence . . .
	. . . if the evidence for clinical utility or clinical validity is insufficient in quantity or quality to support conclusions or make a recommendation.
	Consider the importance of each contextual factor and its magnitude or finding.
	Determine whether the recommendation should be Insufficient (neutral), Insufficient (encouraging), or Insufficient (discouraging).
	Provide information on key information gaps to drive a research agenda.

These clinical practice recommendations are based on the evidence report, the EWG's assessment of the magnitude of net benefit and the certainty of evidence, and consideration of other clinical and contextual issues. Although not clearly stated in the methods, the review of evidence in EGAPP seem to be based on systematic, narrative reviews of the evidence.

As mentioned before in the testing paragraph, Veenstra et al (28) claim that the uncertainty in the evidence around GBT and GBIs, warrants assessing the risk-benefit ratio by using decision-analytic modelling, which will make the assessment more formal and transparent. Based on the amount of certainty of the evidence and the assessment of the risk-benefit profile a recommendation matrix can be developed (26; 31) (see table 4).

Table 7 Risk benefit decision matrix (taken from Roth et al. (31))

		Uncertainty		
		High	*Moderate*	*Low*
Risk-benefit	*Favorable*	Use with evidence development	Consider use in clinical practice	Appropriate for use in clinical practice
	Neutral	Do not use, conduct additional research	Use with evidence development	Consider use in clinical practice
	Unfavorable	Do not use, conduct additional research	Do not use	Do not use

Concluding remarks Synthesis

The synthesis of available evidence of GBTs is different from other new health technologies in that often no direct evidence is available on the effectiveness of the GBT, and assessments have to include indirect evidence. There is guidance and best practice available on medical testing, recently summarized by Bossuyt (25). EGAPP has adapted the USPSTF methodology for rating and grading the available evidence for genome-based applications. A similar issue in assessing and synthesizing the evidence on GBTs compared to other new health technologies is the level of uncertainty of the evidence , due to designs that are not fit to answer questions of efficacy and effectiveness, and lack of studies on clinical validity and utility. Different methodologies can be used (systematic reviews, decision analytic, Bayesian approach). Here the key question is whether the different methods lead to different decisions, and the key issue is the feasibility to produce timely and comprehensive reviews for public health practitioners, and policy makers.

The different methodologies lead to different systems for practice recommendations. What is key, though, is that these systems for practice recommendations are understandable for policy makers and health care professionals. This is not always the case (31).

Dissemination of the synthesized information to the appropriate persons who use or make decisions concerning the use of health technologies and implementation

The objective with both dissemination and implementation is that the information provided by an assessment of a GBT reaches and is used by the target group, health care professionals and policy makers in public health. There is limited research available focusing in particular on policymakers, but factors such as interactions between researchers and health care policy makers and the timing/timeliness of information appear to increase the prospects for research use among policymakers (32;33). In these publications there is no differentiation between dissemination and implementation.

The factor of interaction between researcher and decision maker is present in two visions that emerged from Lehoux et al.'s study (33) into rationales behind dissemination strategies of 6 Canadian HTA agencies. One of these visions is described as pragmatic HTA. In this vision, the purpose of HTA dissemination is certainly to influence evidence-based decision making, but realising that HTA is only one player (and a rather small one) in the arena where policies and decisions are shaped, and that its influence therefore is limited. This vision implies that even though decision- and policy makers remain the most important audiences for HTA, the likelihood that the ultimate outcome will be entirely driven by evidence is limited (33).

An interesting option for agencies adhering to the pragmatic HTA vision, these authors state, would be to cooperate with a third-party organization to disseminate their results, because most active dissemination methods are time consuming and require a certain level of communication skills on the part of evaluators (33). Lomas (34) advocates human intermediaries between the worlds of research and health care practice (knowledge brokers) and supporting infrastructure (knowledge brokering agencies and resources).

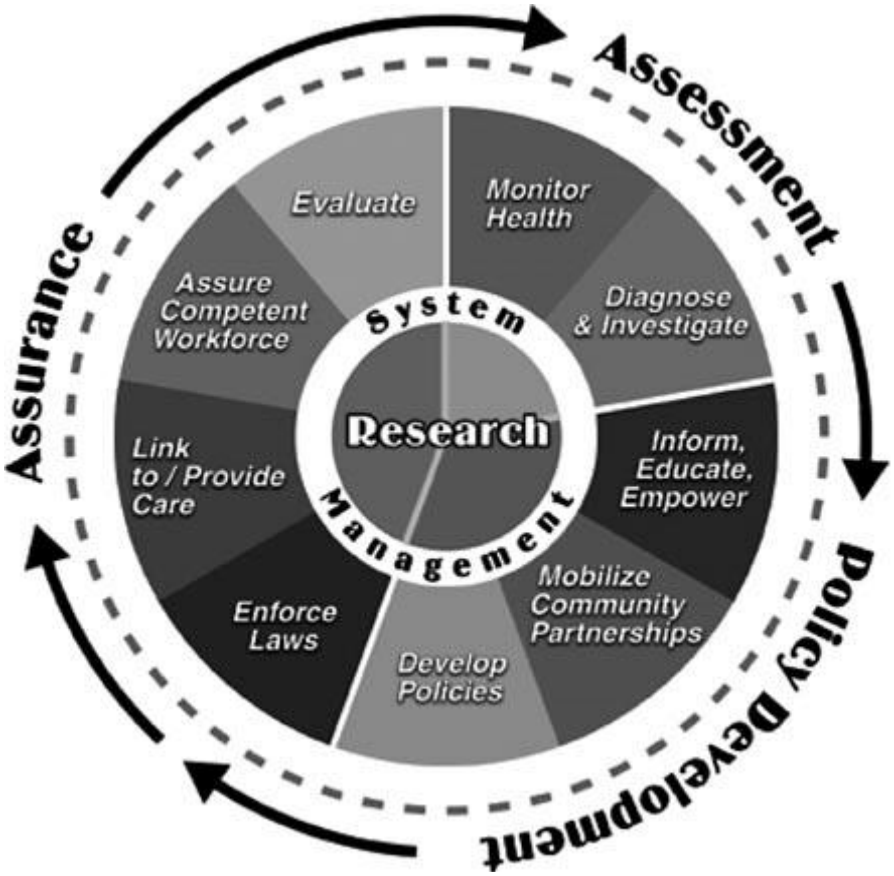
In Lehoux et al.'s (33) other vision, called participatory HTA, users participate in the process of gathering and analysing evidence. The authors explain that in an extreme form of this vision, with a very high level of user involvement during the assessment process, dissemination would become redundant, as the capacity to use HTA would in some way be internalized by a large set of collaborating institutional partners. This vision implies thus a different way of producing HTA, and builds capacity building into the evaluation process. Early and constructiveTA accommodates the involvement of stakeholders during the assessment process.

Dissemination and implementation is also more effective when it is closer to the level of the end-user. For that purpose, it might be beneficial for the translational process of GBT and GBI to implement *localized HTA or mini-HTA*. Mini-HTA is a flexible and dynamic decision support tool, based on the health technology assessment philosophy, adaptable to local conditions and the current requirements of decision-makers – which means that it can relatively easily be incorporated into local and regional budget and planning processes (35). It is originally developed for the hospital setting, but can be adapted to other contexts. Carrying out assessment at the local level, i.e. there where the GBT will be used, will enhance the implementation of the recommendations.

A step that needs to be taken, regardless of which perspective on dissemination or implementation you take is that of improving scientific literacy amongst policy makers, professionals and patients (32). The concept of HTA is quite new in the field of public health. The reason is that the majority of HTA products has focused on clinically oriented interventions, albeit adopting a population perspective (36). So there is a need to educate health care professionals in public health in what Busse et al. (37) call 'best practice activities in the health care sector'. These activities include methods involved and used in HTA, by the Cochrane Collaboration, the EBM movement, and those involved in the creation of clinical practice guidelines (CPG). All activities are characterized by a systematic and structured way of answering

questions by evaluating and synthesizing available evidence (37). In the perspective of this project, early and constructive or early HTA, and mini-HTA should be included.

3. The Guidance (following the wheel per section of the wheel)



ACTION PRINCIPLES (10 Essential Pubic Health Tasks):

Assessment (Tasks 1-2: all three pillars)

3.1 Monitor Health

This function corresponds to the surveillance function in public health which can be viewed as (broad interpretation) applying to two objects:

a) health problems:

- this corresponds to health needs assessment (epidemiology, burden of disease, users' perception of needs…)
- guidance can help define what information should be collected and how…

b) emerging technologies:

- this can be organised as a form of horizon scanning to scan for potential solutions/options to address health needs…
- guidance on how to organise knowledge transfer from basic research and horizon scanning ….

3.1.1 Recommendations

a) Surveillance of health problems:

R1: Surveillance of health problems in the population and subsequent prioritisation of diseases* requires that coded data on the health status of individuals are made available to public health staff responsible for the surveillance (at public health institutes).

Reliable systematic data collection for monitoring purposes can happen through registries and databases that retrieve relevant data from electronic patient records (EPR) using eHealth solutions (data codification, user authentification etc).

Moreover, information on genetic factors (genetic/genomic test results) could be included. This way, registries can provide the tools to support the various types of research belonging to the assessment section of the wheel as outlined by Beskow et al. (epidemiological research, investigate impact of environmental/genetic risk factors …).

Priority should be given to insure the adequate Permanent Research infrastructure and personnel which will allow for the systematic collection of information needed to allow for Health assessment.

R2: Technological aspects of health care systems are rapidly evolving. Investments to implement and keep up with the most adequate ICT solutions including adequate security measures are necessary to allow for qualitative monitoring.

R3: When assessing benefits and risks of health status monitoring and electronic data sharing at national level, possible additional risks from collecting genetic/genomic data should be investigated.

b) Surveillance of health technology applications (= Horizon Scanning)

R1: Clinical utility of GBT/I applications should be evaluated and can only be assessed by monitoring once the GBT/Is are used.

R2: Health Needs Assessment could inform the priorities of Horizon Scanning Systems and HTA agencies for identifying and assessing significant GBT and GBIs in European Union member states

3.1.2 Best Practices

a) Surveillance of health problems

BP1: Genomic/health data collection, storage and their use (analysis) should be done conform existing (inter)national laws w.r.t. data protection, privacy, confidentiality of the data. The OECD 2009 Guidelines on human biobanks and genetic databases should be followed where appropriate. (* OECD2009, Guidelines on Human Biobanks and Genetic Research Databases) * COM(2007) 228: Communication from the Commission to the European Parliament and the Council on Promoting Data Protection by Privacy Enhancing Technology (PETs)

BP2: Appropriate informed consent procedures are necessary (➔autonomy). Consents could –after careful analysis, depending on the context and collected data- evolve towards the opt-out type of consent, which considerably diminishes administrative burden as well as costs.

BP3: When developing national ICT solutions for EPR keeping as well as for registries, (semantic) interoperability* of different systems (abroad*) should be taken into consideration. European recommendations in the e-Health field, including proposed quality criteria, that result from EU projects investigating this, should be adopted by the MS.

(* Recommendation on cross-border interoperability of electronic health records systems, 2008/594/EC)

b) Surveillance of health technology

BP1: Economical criteria should be kept in mind when horizon scanning for useful emerging technology applications.

3.2 Diagnose and Investigate

In line with the above surveillance functions this assessment function can both (broad interpretation):

a) confirm and document the importance of health problems, and

b) evaluate whether the emerging technologies can be effectively/efficiently used and are appropriate to counter/solve health problems.

- The latter (b) thus combines the previous information and relies on (adapted) health technology assessment, including the ELSi (Ethical, Legal, and Social Implications) and organisational components. The ELSi and organisational components should lead to the definition of the optimal conditions of use for the emerging genome-based technologies and information (including safeguards). If projections are made of the likely consequences of use, a kind of a priori health impact assessment can be realised.

Guidance on adapted health technology assessment, a priori health impact assessment …

3.2.1 Recommendations

<u>Diagnose and investigate health problems // health technology applications:</u>

R1: Health monitoring should allow for identification and <u>prioritisation of the diseases (in terms of public health significance)</u> for which more in depth analysis could show the use of GBT/I is worthwile and responsible. Among the criteria: burden of disease, availability of prevention methods/ medical care, important, genomic, including infectious with genomic relevance…

R2: To explicitly link the result of a positive early assessment of GBT and GBIs by an Early Alert/Horizon Scanning System to a potential reduction of the burden of

disease at the population level, and thus inform policy makers on the need to selectively promote the diffusion of particular GBT and GBIs.

R3: Develop results criteria to evaluate whether systems have identified and assessed the most important GBT and GBIs in terms of improved health outcomes.R4: Investigate clinical utility of GBT/I applications after data collection : what are the outcomes if there is a diagnosis (=after application of GBT/I)?

3.2.2 Best Practices

BP1: Care should be taken to ensure the fair allocation of public resources to those who need it most.

.

BP2: Societal impact should also be taken into account.

BP3: A historical example is the development and assessment of new thrombolytic agents in conjunction with Health Needs Assessment in England and Wales as carried out by the NHSC in the UK.

* COM(2007) 228: Communication from the Commission to the European Parliament and the Council on Promoting Data Protection by Privacy Enhancing Technology (PETs)

3.3 Inform, Educate, Empower

- Knowledge transfer to macro, meso and micro level decision-makers to empower them with respect to decisions based on information regarding health problems and needs, adequacy of emerging technologies as solutions, limitations of technologies, gaps in the knowledge base, ELSI issues…
- Guidance on knowledge transfer strategies, on information repositories, on information needs, on optimal presentation of information (on risks for instance), on issues of health literacy….

3.3.1 Recommendations / Best practices

1: There is a need to properly educate and inform all stakeholders in public health genomics including the general public, users (including patients), health care professionals and public health officials and workers regarding GBIs.

1a: General public and "users": (patients and family)

Clear and appropriate information regarding the potential uses, benefits, and harms of GBT/I applications must be made available through different outlets: pamphlets, the Internet and from health care workers. This information should reflect the degree of involvement of the user; for instance an already affected user may need additional information regarding possible treatments and care, whereas for the healthy user information on prevention measures may be more useful. This will allow for an ethical and responsible intervention, by supporting the informed consent process and respecting the autonomy of the patients, as well as increase individual and population acceptance of interventions.

1b: Health care professionals:

Various studies have shown that the majority of non-geneticist physicians are not equipped or prepared to adequately deal with genomic or genetic testing and lack adequate knowledge and training to provide qualitative (genetic) counselling. Those health care professionals most appropriate to deal with GBIs (geneticists, general practitioners, oncologist, paediatricians?) should be identified and appropriate education should be administered in medical education (including continued medical education). Indeed, continued education may be particularly important during the early stages of introduction of GBT/Is since presently practicing physicians who have already completed their formal medical education will also need to be educated.

Education should also reinforce issues of duties in terms of confidentiality privacy protection, but also respect for autonomy besides counselling related principles such as non-directive, etc.

1c: Public health officials and workers:

Improve health and scientific literacy by means of continuous education programmes.

2: To enable patients and consumers to make informed decisions about the benefits and risk of GBT and GBIs, while taking into account values and preferences, by means of decision support tools, including shared-decision making and concordance frameworks and multi-criteria decision analysis based tools (e.g. Annalisa and the Analytical Hierarchy Process).

BP1: Public health Institutes and HTA agencies should make public reports with their conclusions for interested parties such as the general public and policy makers.

3.4 Mobilize Community Partnerships

- This task can be viewed as a major role for the National Task Forces (bringing together genetics, public health, policy makers….) and is an extremely important input for policy development
- Guidance on new business models (health technology life cycle etc.) as well as guidance on stakeholder involvement, setting up of sustainable infrastructure for National Task Forces, ….

3.4.1 Recommendations

a) The National Task Forces could identify relevant stakeholders to be involved in defining the policy question regarding the GBT and GBIs, and the consequent assessment of GBT and GBIs at different stages in their life-cycle.

BP: It is important to include ALL possible stakeholders in this process.

Involvement of patients but also of society as a whole is a prerequisite for the responsible introduction of new tests, this includes decisions on appropriate use of resources.

3.5 Develop Policies

- A broad interpretation can encompass under policies: laws and regulations, governmental policies, professional guidelines, quality assurance measures, different types of safeguards …
- Guidance on existing forms of regulation, development and maintenance of specific forms of regulation, harmonisation issues, …

3.5.1 Recommendations

R1: National Task Forces and national HTA agencies could collaborate to device strategies to effectively and efficiently inform policy makers on the introduction and iterative assessment of GBT and GBIs.

R2: Discrimination and stigmatisation should be avoided or minimized, esp for insurance and work ; take into account when developing policies.

3.5.2 Best practices

BP1: Follow OECD guidelines and European recommendations at macro level when developing policies.

BP2: Follow ISO guidelines at micro level.

Assurance (Tasks 6-9: all three pillars)

3.6 Enforce Laws

- A broad interpretation is needed to encompass the application of the range of policies considered (e.g. Health in All Policies, …).
- Guidance on mechanisms to disseminate and to enforce laws, regulations, policies, quality assurance measures (accreditation, EQA …), means to protect and involve users (informed consent tools, …)

3.6.1 Recommendations

3.6.2 Best practices

BP1 Compliance with specific guidelines, and existing regulations at the national or European level is necessary for the responsible implementation of new tests. This encompasses the following:

1. The actual GBT/Is and their use in public health programmes should meet the applicable legal standards. (OECD)

2. Personal genome information should be subject to the privacy protection and security in accordance with applicable law. (OECD)

3. Advertising, promotional and technical claims for GBTI should accurately describe the characteristics and limitations of the tests offered. (OECD)

3.7 Link to / Provide Care

Two aspects can be envisioned (broad interpretation) relating respectively to the practice and to the organisation of care:

a) Knowledge transfer to clinicians and application of practice guidelines

b) "Translation" of new practices into the required organisational changes (e.g. redefining roles and responsibilities of various professionals…) and implementation of these changes (e.g. planning adequate resources, negotiations…)

3.7.1 Recommendations

Horizon Scanning

a) Investigate the need for set up of a European Early Alert and Assessment System to identify GBT and GBIs that will have a significant impact on European public health priorities (e.g. cancer and HIV/AIDs, health threats including communicable diseases, obesity) or pose a significant threat to consumer safety. Identified GBT and GBIs can be included in the existing EuroScan database, and be publicly available.

b) In the meantime, it can be stimulated to include genome-based technology and information identified by national Early Alert and Assessment Systems in the available database at EuroScan, and ensure public availability.

Allocation of resources/priority setting

c) The shifting paradigm (P4) does not lead to changes in the type of criteria for priority setting of GBT and GBIs, but increases the need for a focused discussion of how GBT and GBIs relate to other new health technologies, in the process of allocating scarce resources for assessment and coverage. GBT and GBIs need to be

related to other health technologies in terms of which contribute most to reduction of the burden of disease or other relevant public health outcomes.

d) If the notion of personal utility is included in clinical utility as a criterion for priority setting, then this should be operationalized for a GBI, and its weighting in relation to other criteria be made explicit.

Testing

e) Frameworks for evaluating the value of GBT and GBIs need to include a comparator.

f) Further research is needed into development or testing of a framework that is both comprehensive and leads to timely, accessible reports, and takes into account the lack of direct evidence, and uncertainty in the available evidence. Combining the strengths of different frameworks, such as the HTA core for its comprehensiveness, risk-benefit analysis, and Early and constructive TA should be part of this research in order to arrive at a framework that evaluates GBT and GBIs in an early stage and thus ensure an efficient translational process..

g) Early and Constructive TA should be stimulated to develop and translate GBT and GBIs that are based on health needs, and ensure the introduction of GBT and GBIs that provide value for money, and allow a rapid translational process.

Synthesis of evidence and recommendations for practice

h) There is a need for adjustment of existing methodologies for in terms of feasibility of assessing the evidence (concerning the amount of resources

necessary), enabling proactive and timely recommendations that are understandable for health care practitioners and policy makers in public health.

Dissemination and implementation into practice

i) There is a need to improve scientific literacy amongst policy makers, professionals and patients in public health and educate them in what Busse et al. (37) call 'best practice activities in the health care sector', including early HTA (or CTA) and local HTA/hospital based HTA.

3.7.2 Best Practices

Horizon Scanning

BP1: GAPP Finder at U.S. CDC's Office of Public Health Genomics.

A searchable database of genetic tests and genomic applications in transition from research to clinical and public health practice (http://www.hugenavigator.net/GAPPKB/topicStartPage.do).

BP2: EuroScan Database: http://euroscan.org.uk/technologies/public/search?advance-search=on

BP3: EuroScan Toolkit for the identification and assessment of new and emerging health technologies http://euroscan.org.uk/methods/

Synthesis

BP4: The Evaluation of Genetic Tests and Genomic Applications (EGAPP) guidelines for evaluation and recommendation (19).

Early HTA (CTA)

BP5: Early and constructive TA of Genetic profiling for Breast cancer (29;30)

Implementation

BP6: The SBU and Alberta Ambassador program to translate HTA findings into practice by educating health care professionals by means of out-reach visits (42).

Best Practices Provide Care

BP1: The introduction of GBT/I should be informed by ethical, social and legal analyses. Relevant ethical, legal and social standards should be respected and promoted in the use of GBT/I.

BP2: The application of GBT/I for individuals or populations should be preceded by the provision of adequate information (forms of counselling).

BP3: The return of results of GBT/I should be done in a responsible manner. Some results of GBT/I may have to be provided with counselling (particularly for serious conditions and/or unfavourable results.

BP4: If forms of treatment or prevention exist for certain applications, these should be offered with a minimum delay.

BP2b: Information in a non-directive manner; voluntary choice, including in the choice of treatments.

BP5: A multidisciplinary approach in health care w.r.t. GBT/I applications is necessary.

3.8 Assure Competent Workforce

- Training of a range of professionals
- Guidance regarding core competencies and communication skills (link with users' information needs…)

3.8.1 Recommendations

R1: There is a need to improve scientific literacy amongst policy makers, professionals and patients in public health and educate them in what Busse et al. (37) call 'best practice activities in the health care sector', including early HTA (or CTA) and local HTA/hospital based HTA.

R2: HTA agencies should include in their assessments the training needs of professionals, policy makers, patients and consumers in relation to the introduction of the GBT and GBIs

R3: Stimulate localized/mini-HTA as decision support tool at the local level in public health to support decisions about the introduction of GBTs and GBIs

R4: HTA agencies should adopt participatory models of carrying out assessments (33) to enhance dissemination and implementation of information from assessments.

R5: A manpower plan should be established for several scientific disciplines: bio-informatics and – statistics, epidemiology, health services research, health technology assessment and health economics.

3.9 Evaluate

Again, two types of evaluation can be thought of here (broad interpretation):

a) Evaluation of the implementation and use of the emerging genome-based technologies and information (and ensuing interventions) :

- Guidance on a posterior health impact assessment

b) Evaluation of the implementation of different types of policies :

- Guidance on policy evaluation

3.9.1 Recommendations

R1: Applications of GBT or GBI need to be evaluated as much as possible before they are introduced into public health programmes. This includes the technical validity, clinical validity, and clinical utility of the applications. Evaluation should occur, where possible, before GBT/I are introduced into public health programmes but also regularly throughout their use. The latter is especially important given that clinical utility can only be fully studied once the GBT/I has been applied.

R2: Appropriate funding and research infrastructure for the full evaluation of GBT/I applications should be established and maintained.

R3: The posterior evaluation process should also include the study of the ethical, legal and social issues raised by the introduction of GBT/I.

R4: Evaluate at the individual ánd population level.

4. Technical Progress

4.1. Overview of activities for the period covered in the interim report (November 2010 - November 2011)

In this period two draft papers were prepared by members of the working group:

Criteria for responsible introduction of genome-based-technologies and information into public health by Heidi Carmen Howard, Elfriede Swinnen , Pascal Borry, Karla Douw, Hindrik Vondeling, and Jean-Jacques Cassiman

Health Technology Assessment of new health technologies, Promoting the understanding of health technology assessment in the field of Public Health Genomics by Hindrik Vondeling (on behalf of Working Group on Quality Assessment of PHGENII) and David Banta.

Furthermore, a questionnaire was developed with the objective of describing the current state-of-the-art and best practice with regard to identifying and assessing genome-based technologies in European Early Alert and Awareness systems. The survey was developed in cooperation with EuroScan, the international Information Network on New and Emerging Health Technologies.

In this process several skype meetings were arranged to discuss the content and development of these products in the group.

4.2 Involvement in the pilots

The group composed by Brettfeld C., Verhaagen M., Englert S. and Haslberger A.G. prepared a preliminary case study report independently from a specific WP, Public health genomics and nutri(epi)genetics: Exploring the wheel.

4.3. Scientific publications

The draft papers mentioned in paragraph 4.1 will be submitted for publication in a peer-reviewed journal.

4.4 Wiki-PHGEN contributions

The group provided the coordinator group continuously with the requested documents for the wiki.

5. Concluding Remarks

The Quality assurance working group provides both a normative and analytic framework for the responsible introduction of genome-based technologies and information. These frameworks can be used to develop local guidelines that are tailored to country-specific and institutional contexts.

6. References

1. Chou AF, Norris AI, Williamson L, Garcia K, Baysinger J, Mulvihill JJ. Quality assurance in medical and public health genetics services: a systematic review. Am J Med Genet C Semin Med Genet. 2009 Aug 15;151C(3):214-34. Review.

2. Ibarreta D, Matthiesen-Guyader L. Quality assurance in genetic testing: a European perspective. Italian J of Public Health, 2006;3:20-23.

3. Rosenkoetter N, Vondeling H, Blancquaert I, Brand A, Kristensen FB, Mekel OCL. The contribution of health technology assessment, health needs assessment, and health impact assessment to the assessment and translation of technologies in the field of public health genomics. Public Health Genomics. 2011;14(1):43-52. Epub 2010 Oct 8.

4. Health Council of the Netherlands. Screening: between hope and hype. The Hague: Health Council of the Netherlands, 2008; publication no. 2008. http://www.gezondheidsraad.nl/en/publications/screening-between-hope-and-hype-0

5. Pernice I, Schulte in den Bäumen T. Public Health Genomics – Legal Aspects of a European Implementation of Genome-based Knowledge and Technologies in Public Health. Ongoing Working Document of the Legal Working Group Paper on Public Health Genomics & Law. 2007 http://www.phgen.eu/typo3/fileadmin/downloads/023_WG4.pdf

6. Khoury MJ, Bowen MS, Burke W, Coates RJ, Dowling NF, Evans JP, Reyes M, St Pierre J. Current priorities for public health practice in addressing the role of human genomics in improving population health. Am J Prev Med. 2011 Apr;40(4):486-93.

7. Khoury MJ, Gwinn M, Yoon PW, Dowling N, Moore CA, Bradley L. The continuum of translation research in genomic medicine: how can we accelerate the appropriate

integration of human genome discoveries into health care and disease prevention? Genet Med. 2007 Oct;9(10):665-74. Review.

8. Khoury M. GAPPNET. A Stakeholder-driven Collaboration to Translate Genome-based Discoveries into Health Benefits. Powerpoint presentation, 2009.

9. Rogowski WH, Hartz SC, John JH. BMC Health Serv Res. 2008 Sep 24;8:194. Clearing up the hazy road from bench to bedside: a framework for integrating the fourth hurdle into translational medicine.

10. Agency for Healthcare Research and Quality. Methods Guide for Medical Test Reviews. [posted November 2010]. Rockville, MD. Available at: http://effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?pageaction=displayProduct&productID=454

11. EuNetHTA. HTA definition. http://www.eunethta.eu/Public/About_EUnetHTA/HTA/ Accessed: 09.12.2011.

12. Banta D, Luce B. Health care technology and its assessment. An international perspective. Oxford: Oxford University Press, 1993.

13. EuroScan. Toolkit for the identification and assessment of new and emerging health technologies. http://euroscan.org.uk/methods/ Accessed: 09.12.2011.

14. Gwinn M, Grossniklaus DA, Yu W, Melillo S, Wulf A, Flome J, Dotson WD, Khoury MJ. Horizon scanning for new genomic tests. Genet Med. 2011 Feb;13(2):161-5.

15. Noorani HZ, Husereau DR, Boudreau R, Skidmore B. Priority setting for health technology assessments: a systematic review of current practical approaches. Int J Technol Assess Health Care. 2007;23(3):310-5.

16. Andermann A, Blancquaert I, Beauchamp S, Costea I.. Guiding policy decisions for genetic screening: developing a systematic and transparent approach. Public Health Genomics. 2011;14(1):9-16. Epub 2009 Dec 29.

17. Henshall C, Oortwijn W, Stevens A, Granados A, Banta D. Priority setting for health technology assessment. Theoretical considerations and practical approaches. Priority setting Subgroup of the EUR-ASSESS Project. Int J Technol Assess Health Care. 1997 Spring;13(2):144-85

18. Douw K, Vondeling H. Selection of new health technologies for assessment aimed at informing decision making: A survey among horizon scanning systems. Int J Technol Assess Health Care. 2006 Spring;22(2):177-83.

19. Teutsch SM, Bradley LA, Palomaki GE, Haddow JE, Piper M, Calonge N, Dotson WD, Douglas MP, Berg AO; EGAPP Working Group. The Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Initiative: methods of the EGAPP Working Group. Genet Med. 2009 Jan;11(1):3-14.

20. Khoury MJ, McBride CM, Schully SD, Ioannidis JP, Feero WG, Janssens AC, Gwinn M, Simons-Morton DG, Bernhardt JM, Cargill M, Chanock SJ, Church GM, Coates RJ, Collins FS, Croyle RT, Davis BR, Downing GJ, Duross A, Friedman S, Gail MH, Ginsburg GS, Green RC, Greene MH, Greenland P, Gulcher JR, Hsu A, Hudson KL, Kardia SL, Kimmel PL, Lauer MS, Miller AM, Offit K, Ransohoff DF, Roberts JS, Rasooly RS, Stefansson K, Terry SF, Teutsch SM, Trepanier A, Wanke KL, Witte JS, Xu J; Centers for Disease Control and Prevention. The Scientific Foundation for personal genomics: recommendations from a National Institutes of Health-Centers for Disease Control and Prevention multidisciplinary workshop. Genet Med. 2009 Aug;11(8):559-67.

21. Grosse SD, McBride CM, Evans JP, Khoury MJ. Personal utility and genomic information: look before you leap. Genet Med. 2009 Aug;11(8):575-6.

22. Foster MW, Mulvihill JJ, Sharp RR.Evaluating the utility of personal genomic information. Genet Med. 2009 Aug;11(8):570-4.

23. Ferrante di Ruffano L, Hyde CJ, McCaffery KJ, Bossuyt PMM, Deeks JJ How do tests impact on patient health? Development of an explanatory framework. Poster at Conference Evidence Live 2011. http://www.evidence2011.com/posters?t=22

24. Lampe K, Mäkelä M, Garrido MV, Anttila H, Autti-Rämö I, Hicks NJ, Hofmann B, Koivisto J, Kunz R, Kärki P, Malmivaara A, Meiesaar K, Reiman-Möttönen P, Norderhaug I, Pasternack I, Ruano-Ravina A, Räsänen P, Saalasti-Koskinen U, Saarni SI, Walin L, Kristensen FB; European network for Health Technology Assessment (EUnetHTA). The HTA core model: a novel method for producing and reporting health technology assessments. Int J Technol Assess Health Care. 2009 Dec;25 Suppl 2:9-20.

25. Bossuyt PMM. Evidence-based medical testing. Developing evidence-based reimbursement recommendations for tests and markers. Report prepared for the Dutch Health care insurance board. In: Medische tests (beoordeling stand van de wetenschap en praktijk). Dutch health care Insurance board, Diemen, 2011. www.cvz.nl/binaries/live/cvzinternet/.../rpt1101-medische-tests.pdf

26. Khoury MJ, Coates RJ, Evans JP. Evidence-based classification of recommendations on use of genomic tests in clinical practice: dealing with insufficient evidence. Genet Med. 2010 Nov;12(11):680-3.

27. Lijmer JG, Leeflang M, Bossuyt PMM. Proposals for a Phased Evaluation of Medical Tests. Medical Tests-White Paper Series [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US); 2009-. AHRQ Methods for Effective Health Care. 2009 Nov 09.

28. Veenstra DL, Roth JA, Garrison LP Jr, Ramsey SD, Burke W. A formal risk-benefit framework for genomic tests: facilitating the appropriate translation of genomics into clinical practice. Genet Med. 2010 Nov;12(11):686-93.

29. Retel V. Constructive Technology Assessment of Gene Expression Profiling for Breast Cancer. PhD. Thesis University of Twente, 2011. http://doc.utwente.nl/78236/1/thesis_V_Retel.pdf

30. Retel V. Hummel M. et al. Constructive Technology Assessment (CTA) as a tool in Coverage with Evidence Development: The case of the 70-gene prognosis signature for breast cancer diagnostics. ISTAHC journal 2009.

31 Roth JA, Garrison LP Jr, Burke W, Ramsey SD, Carlson R, Veenstra DL. Stakeholder perspectives on a risk-benefit framework for genetic testing. Public Health Genomics. 2011;14(2):59-67. Epub 2010 Apr 20.

32. Mowatt G, Thomson MA, Grimshaw J, Grant A. Implementing early warning messages on emerging health technologies. Int J Technol Assess Health Care. 1998;14(4):663-70.

33. Lehoux P, Denis JL, Tailliez S, Hivon M. Dissemination of health technology assessments: identifying the visions guiding an evolving policy innovation in Canada. J Health Polit Policy Law. 2005;30(4):603-41.

34. Lomas J. The in-between world of knowledge brokering. BMJ. 2007;20:334(7585):129-32.

35. Ehlers L, Vestergaard M, Kidholm K, Bonnevie B, Pedersen PH, Jørgensen T, Jensen MF, Kristensen FB, Kjølby M. Doing mini-health technology assessments in hospitals: a new concept of decision support in health care? Int J Technol Assess Health Care. 2006 Summer;22(3):295-301.

36. Battista R. Health technology assessment and public health: a time for convergence. Editorial. Eur J Public Health (2009) 19 (3): 227.

37. Busse R, Orvain J, Velasco M, Perleth M, Drummond M, Gürtner F, Jørgensen T, Jovell A, Malone J, Rüther A, Wild C. Best practice in undertaking and reporting health technology assessments. Working Group 5 Report. Int J Technol Assess Health Care. 2002;18:361-422.

38. Rashiq S, Barton P, Harstall C, Schopflocher D, Taenzer P; Alberta Ambassador Program Team. The Alberta Ambassador Program: delivering Health Technology Assessment results to rural practitioners. BMC Med Educ. 2006 Mar 31;6:21.