



How is genetic testing evaluated? A systematic review of the literature

Erica Pitini¹ · Corrado De Vito¹ · Carolina Marzuillo¹ · Elvira D'Andrea^{1,2} · Annalisa Rosso¹ · Antonio Federici³ · Emilio Di Maria⁴ · Paolo Villari¹

Received: 22 May 2017 / Revised: 7 November 2017 / Accepted: 19 December 2017
© The Author(s) 2018. This article is published with open access

Abstract

Given the rapid development of genetic tests, an assessment of their benefits, risks, and limitations is crucial for public health practice. We performed a systematic review aimed at identifying and comparing the existing evaluation frameworks for genetic tests. We searched PUBMED, SCOPUS, ISI Web of Knowledge, Google Scholar, Google, and gray literature sources for any documents describing such frameworks. We identified 29 evaluation frameworks published between 2000 and 2017, mostly based on the ACCE Framework ($n = 13$ models), or on the HTA process ($n = 6$), or both ($n = 2$). Others refer to the Wilson and Jungner screening criteria ($n = 3$) or to a mixture of different criteria ($n = 5$). Due to the widespread use of the ACCE Framework, the most frequently used evaluation criteria are analytic and clinical validity, clinical utility and ethical, legal and social implications. Less attention is given to the context of implementation. An economic dimension is always considered, but not in great detail. Consideration of delivery models, organizational aspects, and consumer viewpoint is often lacking. A deeper analysis of such context-related evaluation dimensions may strengthen a comprehensive evaluation of genetic tests and support the decision-making process.

Introduction

The increased availability of genetic tests has made the assessment of their performance crucial for clinical and public health practice. However, the evaluation of genetic tests, especially predictive ones, is not straightforward. The main challenge is the lack of scientific evidence on which to base such evaluations [1]. Generating scientific evidence on genetic tests is made difficult by different factors including their complexity, their rapid development and marketing,

their widespread impact on families and society, and the lack of standardized outcomes for their evaluation [1]. For predictive genetic tests, perhaps the greatest challenge is to perform high quality, randomized control trials to demonstrate that the test confers an improvement in survival or quality of life [2]. Moreover, the lack of evidence on effectiveness affects the evaluation of cost-effectiveness.

Despite this, several frameworks have been proposed for the evaluation of genetic tests, but it is unclear how and in what respect they differ. The importance of a transparent and well-planned evaluation strategy is twofold. On the one hand, it would avoid the uncontrolled implementation of technologies without proven benefits, which can lead to inappropriate management of patients and detrimental effects on patient health, as well as a waste of resources and loss of public confidence in the medical profession. On the other hand, in line with the requirement for public health programs to maximize population health benefits, a reliable evaluation strategy would support the implementation of those currently available tests that have proven effectiveness and cost effectiveness [3].

To guide the appropriate translation of genomics into clinical practice, Italy developed a National Plan for Public Health Genomics, which, to our knowledge, is the first specific policy example of its kind in Europe. It has various

Electronic supplementary material The online version of this article (<https://doi.org/10.1038/s41431-018-0095-5>) contains supplementary material, which is available to authorized users.

✉ Erica Pitini
erica.pitini@uniroma1.it

¹ Department of Public Health and Infectious Diseases, Sapienza University of Rome, Rome, Italy

² Harvard TH Chan School of Public Health, Boston, MA, USA

³ Ministry of Health, Rome, Italy

⁴ Department of Health Sciences, Division of Medical Genetics, University of Genova, Galliera Hospital, Genova, Italy

strategic objectives including the development of a well-planned evaluation strategy for genetic tests [4]. Our systematic review was conducted as part of a project financed by the Italian Ministry of Health to implement this plan and aims to identify and compare the existing evaluation frameworks for genetic tests, taking into account their methodology and evaluation criteria.

Materials and methods

This review was performed according to the Cochrane Handbook for systematic reviews of intervention and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [5, 6].

Selection criteria

We included any document that describes an original evaluation framework for genetic tests, defined as a structured process for the collection of the scientific evidence needed to assess the performance of a genetic test, from the laboratory to clinical practice. We excluded partial evaluation frameworks, defined as those embracing less than three evaluation components (analytic validity, clinical effectiveness, etc.). We limited our search to frameworks specifically created for the evaluation of genetic tests.

Search methods

Two reviewers (EP and CD) searched the bibliographic databases Pubmed, Scopus, ISI Web of Knowledge, Google Scholar, and the world wide web through Google for all English language articles between January 1990 and April 2017. The search terms were grouped as two strings: String A, “genetic testing OR genetic test OR genomic test OR genomic technology OR pharmacogenetic test” AND “evaluation OR assessment OR evaluating OR assessing OR evaluate OR assess” AND “framework OR criteria OR tool OR model OR process OR methods OR evidence based” OR “analytic validity OR clinical validity OR clinical utility”; and String B, “genetic testing OR genetic test OR genomic test OR genomic technologies OR pharmacogenetic test OR public health genomics OR pharmacogenetics OR pharmacogenomics” AND “health technology assessment” (see Supplementary Information for the full electronic search for each database). This search was supplemented by exploring the websites of government agencies and research organizations involved in the evaluation of genetic tests (see Supplementary Information for a list of the respective websites) and by scanning the reference lists of all the relevant articles retrieved. Moreover, experts of the Italian

Network of Public Health Genomics were asked to share the evaluation frameworks they were aware of through a Delphi procedure.

Study selection

The two reviewers (EP and CD) removed duplicates and screened the title and abstract of all retrieved records. Studies that clearly did not meet the eligibility criteria were excluded. Full texts of potentially relevant studies were examined for inclusion in the systematic review and reasons for exclusion recorded. Disagreements were resolved by discussion.

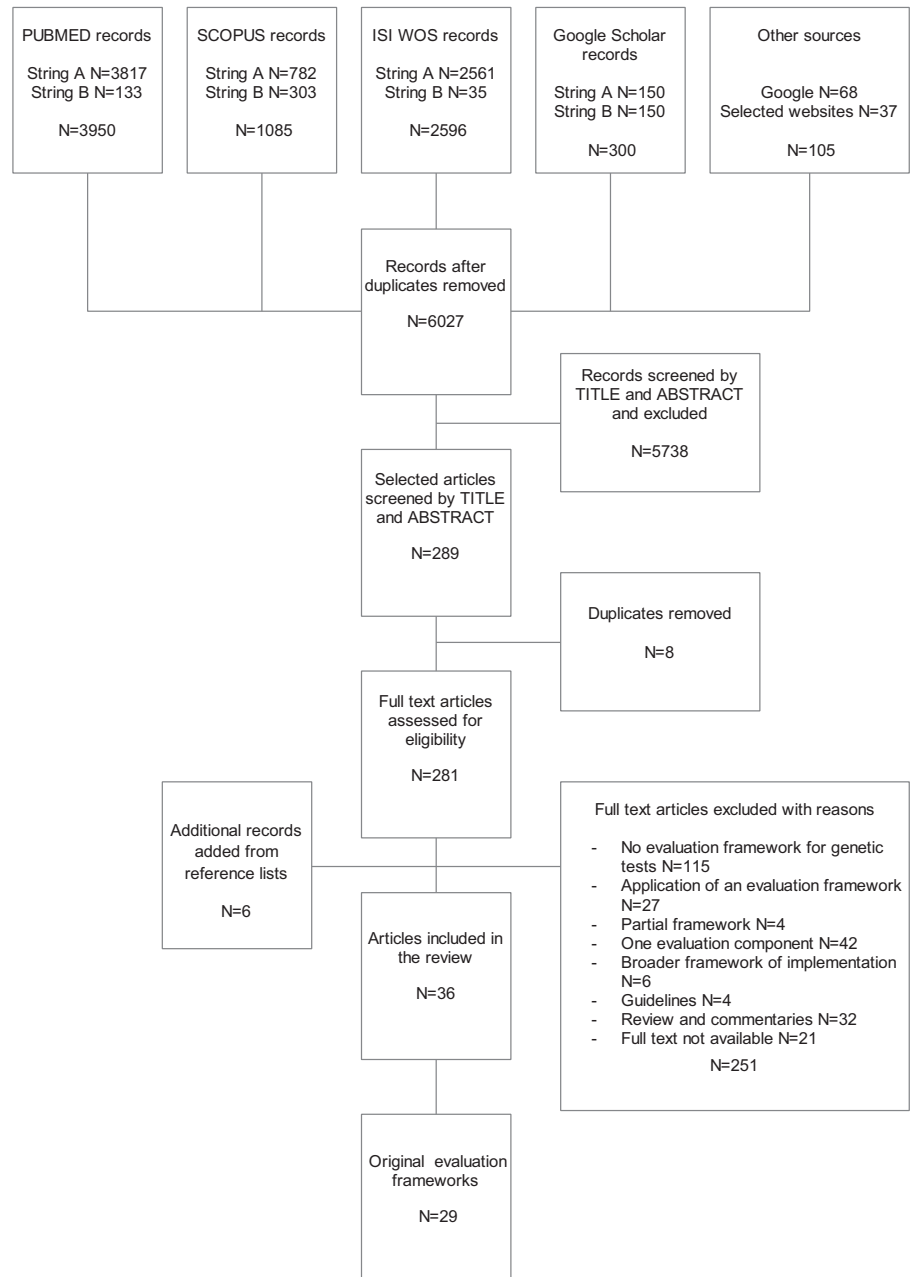
Data collection and analysis

Two reviewers (EP and EDA) extracted the following information about the retrieved frameworks: authors; country; year of publication; reference institution; framework name; type of target test; reference frameworks; methodology (format, sources of evidence, quality of the evidence, grading of recommendations, research priorities); practical application; purpose; primary audience; evaluation components (see Supplementary Information for a definition of each category of information extracted). A narrative synthesis of the evaluation frameworks identified was performed, comparing their general features, evaluation components, and methodological aspects.

Results

Study selection

After removal of duplicates, 6027 records resulted from the initial search (Fig. 1). Screening by title and abstract selected 289 records for full text analysis, from which 30 records were selected. Reasons for exclusion were: documents not describing a framework for the evaluation of genetic tests; appraisals of individual genetic tests using an original framework described elsewhere; partial evaluation frameworks; documents focusing on only one evaluation component; broader frameworks of implementation research not proposing an original evaluation process; guidelines on the evaluation of genetic tests; reviews and commentaries of evaluation frameworks and evaluation criteria for genetic tests; full text not available. Six records were added to the previous 30 from the reference lists of relevant articles retrieved. A total of 36 records were included in the systematic review, describing 29 frameworks for the evaluation of genetic tests (some records describe the same framework) [7–42]. The Delphi procedure did not add new frameworks to those already retrieved.

Fig. 1 PRISMA flow diagram of the review process

Frameworks retrieved

The systematic search identified 29 frameworks from various countries (USA, $n = 12$; Canada, $n = 4$; Europe, $n = 9$; Australia, $n = 2$; international, $n = 2$) published between 2000 and 2017 (Table 1).

The majority are based on the ACCE Framework (whose name derives from the evaluation components used: analytic validity, clinical validity, clinical utility, ethical, legal and social implications) ($n = 13$) [7–24], on the Health Technology Assessment (HTA) process ($n = 6$) [25–31], or both ($n = 2$) [32, 33]. The remaining frameworks refer to the Wilson and Jungner screening criteria ($n = 3$) [34–37],

or to a mixture of preexisting frameworks, which are not necessarily specific for genetic tests, even if the ACCE framework is often included ($n = 5$; Table 1) [38–42].

Seventeen frameworks deal with genetic tests in general [7–19, 22, 26, 27, 32, 36–42]; five refer to genetic susceptibility tests [20, 24, 33–35]; three to pharmacogenetic tests [23, 30, 31]; two to predictive genetic tests (including susceptibility and presymptomatic tests) [25, 28]; one to the new technologies of personalized medicine [29] and another to newborn screening [21] (see Supplementary Information for definitions of types of test) (Table 1). Most of the frameworks pursue a wider aim than simply summarize evidence, for example, support provision and coverage

Table 1 Frameworks used to evaluate genetic tests retrieved by the systematic review process

Framework name	Institution/ organization ^a	Year ^b	Country ^c	Reference frameworks	Target test	Purpose	Primary audience	Practical application	Ref.
ACCE	CDC	2000	USA	ACCE	Genetic tests	Clinical guidance and policy making	Policy makers, health professionals, general public	Website CDC	[7, 43]
NHS UKGTN Gene Dossier	UKGTN	2004	UK	ACCE	Genetic tests	Provision and coverage decisions	Decision makers	Website UKGTN	[8–10, 44]
EGAPP	EGAPP	2005	USA	ACCE	Genetic tests	Clinical guidance and policy making	Clinicians and other key stakeholders	Website EGAPP	[11, 12, 45]
Andalusian	AETSA	2006	Spain	ACCE	Genetic tests	Provision and coverage decisions	Decision makers	Not available	[13]
Rapid ACCE	None	2007	USA	ACCE	Genetic tests	Clinical guidance and policy making	Policy makers, physicians, patients	Genetics in Medicine Journal	[14, 46]
Expanded ACCE	PHGf	2007	UK	ACCE	Genetic tests	Clinical guidance and policy making	Policy makers	Not available	[15]
GFH Indication Criteria	GFH	2008	Germany	ACCE	Genetic tests	Clinical guidance	Referrers, service providers, payers	GFH website (German)/ ESHG website (English)	[16, 47]
Clinical Utility Gene Card	EuroGentest	2010	Europe	ACCE	Genetic tests	Clinical guidance	Clinicians, geneticists, referrers, service providers, payers	Website EuroGentest/ Website European Journal of Human genetics	[17–19, 48, 49]
Complex Diseases	None	2010	UK	ACCE	Susceptibility tests	Clinical guidance and policy making	Policy makers, physicians, general public	Not available	[20]
ACHDNC	ACHDNC	2010	USA	ACCE	Newborn screening	Decision on inclusion in the recommended newborn screening panel	Decision makers	Website ACHDNC	[21, 50]
ECRI	ECRI	2011	USA	ACCE	Genetic tests	Provision decisions	Patients, providers, payers, regulators, test developers	AHRQ website	[22, 51]
CAT	None	2015	USA	ACCE	Pharmacogenetic tests	Evidence summary	Decision makers	Not available	[23]
Clinical Utility Card	MSAC	2016	Australia	ACCE	Susceptibility test	Provision and coverage decisions	Decision makers	MSAC website	[24, 52]
PACNPGT	PACNPGT	2001	Canada	HTA	Predictive tests	Provision and coverage decisions	Decision makers	Not available	[25]
AETMIS HTA	AETMIS	2001	Canada	HTA	Genetic tests	Policy making	Policy makers	Website AETMIS	[26, 27, 53]
MC Master	MC Master University	2003	Canada	HTA	Predictive tests	Provision and coverage decisions	Decision makers	Not available	[28]
HTA Personalized Health Care	None	2012	Spain	HTA	Personalized Health Care	Provision and coverage decisions	Decision makers	Not available	[29]
Codependent Technologies	None	2013	Australia	HTA	Pharmacogenetic tests	Provision and coverage decisions	Decision makers	Not available	[30]
HTA Pharmacogenetics	None	2013	UK	HTA	Pharmacogenetic tests	Clinical guidance and policy making	Policy makers, clinicians	Appraisal itself	[31]

Table 1 (continued)

Framework name	Institution/organization ^a	Year ^b	Country ^c	Reference frameworks	Target test	Purpose	Primary audience	Practical application	Ref.
Hayes GTE	Hayes Inc	2008	USA	HTA-ACCE	Genetic tests	Provision decisions	Policy makers, insurers, hospitals	Website Hayes Inc (with fee)	[32, 54]
HTA Susceptibility Test	Italian Ministry for Education and University	2012	Italy	HTA-ACCE	Susceptibility tests	Clinical guidance and policy making	Policy makers, physicians, citizens, other stakeholders	Appraisal itself	[33]
Cancer Screening	Crossroads 99 group	2001	Canada	Wilson and Jungner principles	Susceptibility tests in screening programs	Policy making	Policy makers	Not available	[34]
Chronic Disease Screening	None	2001	USA	Wilson and Jungner principles	Susceptibility tests in screening programs	Policy making	Policy makers	Not available	[35]
Genetic Screening	None	2011	International	Wilson and Jungner principles	Genetic tests in screening programs	Policy making	Policy makers, other stakeholders	Not available	[36, 37]
GETT	IFCC	2010	International	Various	Genetic tests	Provision decisions	Policy makers, researchers, clinical laboratory scientists and administrators, health care system managers	Website CanGeneTest ^d	[38]
SymFRAME	None	2012	USA	Various	Genetic tests	Clinical guidance and policy making	Policy makers, innovators, health technology appraisers, physicians, patients, regulators	Not available	[39]
Practical Framework	None	2012	USA	Various	Genetic tests	Clinical guidance and policy making	Policy makers, clinicians, laboratorians, researchers, test developers, patients	Not available	[40]
Six-part framework	None	2014	USA	Various	Genetic tests	Provision and coverage decisions	Payers	Not available	[41]
NASEM Framework	NASEM	2017	USA	Various	Genetic tests	Provision decisions	Decision makers	Not available	[42]

ACCE Analytic validity, Clinical utility, Ethical, legal and Social Implications, CDC Centers for Disease Control and Prevention, NHS UKGTV National Health Service UK Genetic Testing Network, EGAPP Evaluation of Genomic Applications in Practice and Prevention, AETSA Agencia de Evaluación de Tecnologías Sanitarias (Andalusian Agency for Health Technology Assessment), PHGfPHG foundation, GFH Deutsche Gesellschaft Für Humangenetik (German Society of Human Genetics), ESHG European Society of Human Genetics, ACHDNC Advisory Committee on Heritable Disorders in Newborns and Children, MSAC Medical Service Advisory Committee (Australian Government, Department of Health), AHRQ Agency for Healthcare Research and Quality, PACNPGT Provincial Advisory Committee on New Predictive Genetic Technologies, AETMS Agence d'évaluation des technologies et des modes d'intervention en santé (Quebec Agency for Health Services Research and Technology Assessment), Hayes GTE Hayes Genetic Test Evaluation, GETT Genetic testing Evidence Tracking Tool, IFCC International Federation of Clinical Chemistry and Laboratory Medicine, NASEM National Academies of Sciences, Engineering, and Medicine

^aInstitution/organization: name of the institution/organization supporting the development of the framework

^bYear: year of development or year of publication

^cCountry: country of the reference institution/organization or of the first author

^dThe authors of the GETT framework declared that it has been used to evaluate five genetic tests but we were not able to find these appraisals

decisions or guide clinical practice; the intended primary audience is mainly represented by decision/policy makers (Table 1).

In addition to the two frameworks created as appraisal tools for individual genetic tests (HTA Pharmacogenetics and HTA Susceptibility Test) [31, 33], 11 frameworks were used to generate reports that are available on the web (all open access, except Hayes GTE reports) (Table 1) [43–54]. The most productive frameworks are the Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Process (11 evidence reports and 10 recommendations), the GFH Card (46 cards), the Clinical Utility Gene Card (about 136 cards), and the NHS UKGTN Gene Dossier (476 dossiers) [44, 45, 47–49].

Evaluation components

The most-represented evaluation components in the retrieved frameworks are analytic validity (included in 93% of retrieved frameworks), clinical validity (96%), clinical utility (100%), economic validity (100%), and ethical, legal, and social implications (ELSI) (76%). The analysis of these components is usually introduced by an overview of the disease and the test under study (86%). Evaluation components frequently missing from the evaluation frameworks are organizational aspects (lacking in 48% of retrieved frameworks), delivery models (73%), and the patient/citizen's point of view (93%) (Table 2).

Analytic validity is the ability of the test to accurately and reliably measure the genotype of interest [55]. It is considered in markedly different ways in 27 retrieved frameworks (Table 2) [7–28, 30–34, 36–42]. It is most frequently addressed in terms of sensitivity and specificity, but assay robustness and quality assurance, including internal and external control programs, are often considered (e.g., ACCE, EGAPP) [7, 11, 12]; some frameworks (e.g., Expanded ACCE, SynFRAME) extend the concept using more criteria [15, 39].

Clinical validity is the ability of the test to accurately and reliably detect or predict a clinical condition [55]. It is considered in 28 retrieved frameworks (Table 2) [7–28, 30–42]. The majority identify clinical validity as test performance and measure it in terms of sensitivity, specificity, and positive and negative predictive value. Other frameworks (e.g., Expanded ACCE, Complex Diseases) preface the evaluation of the performance of the test with explicit evaluation of the scientific validity, that is, the evidence of gene–disease association, which is usually expressed as an odds ratio or relative risk [15, 20].

Clinical utility, in its narrowest sense, compares the risks and benefits of testing and provides evidence of clinical usefulness for the integrated package of care in terms of measurable health outcomes [56]. It is considered in all 29

Table 2 Evaluation components and methodological aspects considered in the retrieved evaluation frameworks

	<i>N</i> (29)	%	Reference
Evaluation components			
Overview disease/test under study	25	86	[7–22, 24, 25, 27, 29–34, 36–42]
Analytic validity	27	93	[7–28, 30–34, 36–42]
Clinical validity	28	96	[7–28, 30–42]
Clinical utility	29	100	[7–42]
Ethical, legal, and social implications	22	76	[7–15, 20–22, 25–29, 32–40, 42]
Delivery models	8	27	[13, 24, 29, 30, 34–38]
Organizational aspects	15	52	[7–10, 13, 14, 24–30, 32, 34–38]
Economic evaluation ^a	29	100	[7–42]
Patient/citizen's point of view ^b	2	7	[15, 33]
Methodological aspects			
Format			
Key questions	12	41	[7, 11–15, 20–23, 32, 41, 42]
Card	5	17	[8–10, 16–19, 24, 25]
Checklist	2	7	[30, 38]
Set of principle/methodological guidance	10	34	[26–29, 31, 33–37, 39, 40]
Evidence collection and evaluation			
Source of evidence ^c	13	45	[7, 11–15, 21, 22, 25, 30, 31, 33, 39, 42]
Quality of the evidence ^d	12	41	[11–14, 21, 22, 25, 30, 31, 33, 35, 39, 42]
Evidence gaps/research priorities	12	41	[7, 11, 12, 14, 21, 23, 25, 28, 31, 34, 38, 40, 42]
Recommendations	5	17	[11, 12, 21, 28, 32, 42]

Note: where it was unclear whether an evaluation component was present, it was considered missing and the methodological aspect unfulfilled

^aEconomic evaluation: the evaluation component was considered as present in the evaluation framework even if only a cost-analysis, and not a complete economic evaluation, was mentioned

^bPatient/citizen's point of view: the evaluation component was counted only if a direct reference to patient consultation was made

^cSource of evidence: the methodological aspect was considered fulfilled even if the topic of the source of evidence was given only a brief mention without further details

^dEvaluation of quality of the evidence: the methodological aspect was considered fulfilled even if the topic of the quality of evidence was only mentioned briefly without details of the criteria used to assess the quality

frameworks retrieved, with a certain heterogeneity (Table 2) [7–42]. Some frameworks (e.g., HTA Personalized Health Care, HTA Pharmacogenetics) embrace the narrow definition of clinical utility and consider only efficacy, effectiveness, and safety [29, 31]. Others extend the concept and

include considerations of aspects otherwise evaluated independently from clinical utility, such as organizational aspects, cost-effectiveness analysis, and ELSI (e.g., ACCE, Expanded ACCE, ACHDNC) [7, 15, 21]. The broadening of the perception of benefits reaches its greatest extent in the concept of personal utility, adopted in some frameworks (e.g., Complex Diseases, ACHDNC), that is, the full range of personal effects that the test may have on patients, such as improved understanding of the disease, enabling reproductive choices or risk-reducing behaviors [20, 21, 57].

The economic evaluation of genetic tests involves the comparative analysis of both the costs and consequences of the various tests under study [58]. All 29 frameworks retrieved consider the economic dimension, but not always in great detail (Table 2) [7–42]. Many models address the cost-effectiveness of the test under study only in the most general terms (e.g., ACCE), and only a few involve precise quantitative and qualitative evaluations of cost-effectiveness and cost-utility evidence (e.g., Codependent Technologies, SynFrame) [7, 30, 39]; other frameworks consider only the financial aspects, either in terms of the cost of the intervention or related savings (e.g., NHS UKGTN Gene Dossier, PACNPGT) [8–10, 25].

The ELSI evaluation component is concerned with the moral value that society confers on the proposed interventions, the specific related legal norms and the impact on the social life of the patient and his or her family [55]. They are considered in 22 retrieved frameworks and analyzed independently (e.g., ACCE, Andalusian) or integrated into other components of evaluation, such as clinical utility, as psychosocial outcomes of testing (e.g., EGAPP, AETMIS HTA) (Table 2) [7–15, 20–22, 25–29, 32–40, 42].

We defined a delivery model for the provision of genetic tests as the broad context in which genetic tests are offered to individuals and families with or at risk of genetic disorders. It includes the health care programs (any type of health intervention preceding and following a genetic test), the clinical pathways (the patient flow through different professionals during the testing process), and the level of care (e.g., primary or specialist care level) in which the test is delivered [59]. Although a complete description of the delivery models is lacking in all frameworks retrieved, we ascribe it to eight retrieved frameworks (Table 2): the three screening frameworks, as they include the concept of health care program [34–37], and five other frameworks, which mentioned some of the elements, albeit not in detail (Table 2) [13, 24, 29, 30, 38].

Organizational aspects include the human, material, and economic resources needed to implement the genetic program as well as the consequences of the implementation on the organizations involved and the whole health care system. Although they do not include a thorough feasibility analysis, 15 retrieved frameworks attempt to estimate the

resources required to start up and maintain a particular genetic testing service (Table 2) [7–10, 13, 14, 24–30, 32, 34–38].

The perspective of patients provides experiential evidence that can be used in the evaluation process [60]. Only two of the retrieved frameworks evaluate the direct experience of the patient and other affected individuals, for example, by their participation in surveys (Table 2) [15, 33]. One of these (Expanded ACCE) includes the patient perspective as part of the clinical utility component, whereas the other (HTA Susceptibility Test) assigns it its own dedicated section.

Methodological aspects

The most frequently used formats are the key questions format (12 frameworks) [7, 11–15, 20–23, 32, 41, 42], the card format (five frameworks) [8–10, 16–19, 24, 25], and the checklist format (two frameworks) [30, 38]. The other frameworks have a less structured format and resemble general manuals (Table 2) [26–29, 31, 33–37, 39, 40].

With respect to the process of evidence review, 13 frameworks provide some indication, albeit in scant detail, of the sources referred to [7, 11–15, 21, 22, 25, 30, 31, 33, 39, 42]; 12 frameworks provide an evaluation of the quality of evidence, but the criteria adopted are not always stated clearly [11–14, 21, 22, 25, 30, 31, 33, 35, 39, 42]; and finally 12 frameworks attempt to deal with evidence gaps through the formulation of research priorities (Table 2) [7, 11, 12, 14, 21, 23, 25, 28, 31, 34, 38, 40, 42].

Only five of the retrieved frameworks provide, or at least suggest, criteria for making recommendations based on the evidence collected [11, 12, 21, 28, 32, 42]. The most frequently used criteria are the magnitude of the net benefit and the level of certainty of the evidence.

Discussion

Our review identified three main approaches to the evaluation of genetic testing: the ACCE model, the HTA process, and the Wilson and Jungner screening criteria. The most popular is the ACCE model, developed in 2000 by the US Centers for Disease Control and Prevention [7]. In 2004, it was further developed to become EGAPP initiative, to make recommendations for clinical and public health practice [11, 12]. The UK Genetic Testing Network and the Andalusian Agency for Health Technology Assessment re-elaborated the ACCE model to guide the introduction of new genetic tests into their public health system, creating the 2004 NHS UKGTN Gene Dossier and the 2006 Andalusian Framework, respectively [8–10, 13]. In 2007,

the ACCE model was reworked again: an expanded version of ACCE, supported by the PHG Foundation, added health quality measures to the evaluation process, whereas a more streamlined version shortened the systematic review process for emerging genetic tests [14, 15]. In 2010, the ACCE model was applied to specific types of genetic test through the Complex Disease Framework and the ACHDNC Newborn Screening Framework of the Advisory Committee on Heritable Disorders in Newborns and Children [20, 21]. The ACCE model also inspired two related European frameworks, the 2008 GFH Indication Criteria of the German Society of Human Genetics and the 2010 Clinical Utility Gene Card of EuroGentest [16–19], the latter of which in turn inspired the 2017 Australian Clinical Utility Card [24]. In 2011, the ECRI Institute used the EGAPP process to develop a set of analytical frameworks for different testing scenarios and stakeholder perspectives [22]. Finally, the 2015 Companion tests Assessment Tool (CAT) used the ACCE model as a filter mechanism to determine which tests, in which specific areas, required evaluation [23]. Our results show that the ACCE framework is the main conceptual frame for the evaluation of genetic tests and the most used in practice, as it inspired very productive frameworks in terms of evidence reports such as the EGAPP, the Clinical Utility Gene Card, and the NHS UKGTN Gene Dossier. Some attempts have been made to merge the ACCE and the HTA model, for example, the 2009 framework for genetic tests used by the private American company Hayes and the 2012 framework for susceptibility tests financed by the Italian Ministry of Education, Universities and Research [32, 33].

Due to the widespread use of the ACCE framework, the most frequently employed evaluation components are analytic validity, clinical validity, clinical utility, and ELSI. Although these evaluation components clearly address the technical value of a genetic test, less attention is given to the wider context. Thus, although the clinical context is given sufficient consideration in some cases, in particular the NHS UKGTN evaluation process, whose testing criteria define the appropriate clinical situations for use of a given test [60], the broader context for implementation of a genetic test is often disregarded. In fact, even where an economic evaluation is performed, it is usually rather superficial; similarly, the analysis of delivery models and organizational aspects, when presented, is usually not well structured. These context-related evaluation components are more often considered by the HTA-based evaluation frameworks. The direct experience of patients is almost totally neglected. Nevertheless, being patients the direct beneficiaries of a genetic technology, their perspective could help understanding its value [61]. Finally, the criteria for making recommendations on the clinical implementation of tests are rarely explored.

Since decision makers are the main audience of the evaluation process, the lack of attention to the context-related evaluation components (delivery models, economic evaluation, and organizational aspects) and to the recommendation-making process are arguably the main limitations of the retrieved frameworks. The analysis of the context of implementation, peculiar to the HTA, is critical for securing an efficient and equitable allocation of health care resources and services. An EU-funded research project named HIScreenDiag [62], which closed in 2011, aimed to assess genetic tests using the HTA methodology of the European Network for Health Technology Assessment, which includes a detailed analysis of the economic, organizational, and delivery aspects [63]. Moreover, the adoption of an evidence grading system such as Grading of Recommendations Assessment, Development and Evaluation (GRADE), which scores the strength of recommendations after taking into consideration aspects like patient values and use of resources, would help move the evaluation process from evidence to implementation, and would make the process more comprehensive [64]. Finally, because these frameworks were mainly developed to address single-gene testing, we might consider how appropriate they are for tests based on next-generation sequencing (NGS). However, frameworks that have been adapted for NGS, such as the NHS UKGTN Gene Dossier and the Clinical Utility Gene Card, have proved effective [65, 66].

In contrast to our systematic review, the majority of reviews in the literature on evaluation frameworks for genetic tests have a narrative structure. The only three systematic reviews we retrieved are described as methods for the construction of an evaluation framework (ECRI, SynFRAME, Practical Framework), so their methodology is not strictly reported [22, 39, 40].

One limitation of our work might be a failure to retrieve some of the studies published in the gray literature. To maximize the sensitivity of the search, we used broad search terms, but these yielded results with low specificity; however, we corrected this during the selection process. Moreover, the comparison between the retrieved frameworks could have been affected by the fact that not all frameworks clearly defined their evaluation components, especially with respect to delivery models and organizational aspects.

In conclusion, the ACCE model proves to be a base for the technical appraisal of genetic tests. However, this model is not completely satisfying. We suggest the adoption of a broader HTA approach, including the assessment of the context-related evaluation dimensions (delivery models, economic evaluation, and organizational aspects). This approach would maximize population health benefits, facilitate decision-making and address the main challenges

of the implementation of genetic tests, particularly in universal health care systems, where economic sustainability is a major issue.

Acknowledgements This work is part of the Italian project “Definizione e promozione di programmi per il sostegno all’attuazione del Piano d’Intesa del 13/3/13 recante Linee di indirizzo su La Genomica in Sanità Pubblica (Definition and promotion of programs to support the implementation of the Guidelines on Genomics in Public Health)”, funded by the Italian Ministry of Health. This work is also partially supported by the project “Personalized pREvention of Chronic Diseases consortium (PRECeDI)” funded by the European Union’s Horizon 2020 research and innovation programme MSCA-RISE-2014 (Marie Skłodowska-Curie Research and Innovation Staff Exchange), under the grant agreement N°645740.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, and provide a link to the Creative Commons license. You do not have permission under this license to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. 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;12:680–3.
2. Marzuillo C, De Vito C, D’Andrea E, Rosso A, Villari P. Predictive genetic testing for complex diseases: a public health perspective. *QJM*. 2014;107:93–97.
3. Khoury MJ, Bowen MS, Burke W, et al. Current priorities for public health practice in addressing the role of human genomics in improving population health. *Am J Prev Med*. 2011;40:486–93.
4. Boccia S, Federici A, Colotto M, Villari P. Implementation of Italian guidelines on public health genomics in Italy: a challenging policy of the NHS. *Epidemiol Prev*. 2014;38:29–34.
5. Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions*, version 5.1.0. The Cochrane Collaboration, 2011.
6. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol*. 2009;62:e1–34. <https://doi.org/10.1016/j.jclinepi.2009.06.006>.
7. Haddow JE, Palomaki GE. ACCE: a model process for evaluating data on emerging genetic tests. In: Khoury M, Little J, Burke W editors. *Human genome epidemiology: a scientific foundation for using genetic information to improve health and prevent disease*. Oxford, UK: Oxford University Press; 2004. p. 217–33.
8. UK Genetic Testing Network. Testing criteria for molecular genetic tests. 2005. Available at http://ukgtn.nhs.uk/fileadmin/uploads/ukgtn/Documents/Resources/Library/Policies_Procedures/MKTesting%20Criteria%20Paper%202005.pdf. Accessed 26 Apr 2017.
9. Kroese M, Zimmern RL, Farndon P, Stewart F, Whittaker J. How can genetic tests be evaluated for clinical use? Experience of the UK Genetic Testing Network. *Eur J Hum Genet*. 2007;9:917–21.
10. UK Genetic Testing Network. First report of the UKGTN. Supporting genetic testing in the NHS. London, UK: UK Genetic Testing Network; 2008.
11. Teutsch SM, Bradley LA, Palomaki GE, et al. The evaluation of genomic applications in practice and prevention (EGAPP) initiative: methods of the EGAPP Working Group. *Genet Med*. 2009;1:3–14.
12. Veenstra DL, Piper M, Haddow JE, et al. Improving the efficiency and relevance of evidence-based recommendations in the era of whole-genome sequencing: an EGAPP methods update. *Genet Med*. 2013;1:14–24.
13. Márquez Calderón S, Briones Pérez, de la Blanca E. Framework for the assessment of genetic testing in the Andalusian Public Health System. Seville, Andalusia: Andalusian Agency for Health Technology Assessment; 2006.
14. Gudgeon JM, McClain MR, Palomaki GE, Williams MS. Rapid ACCE: experience with a rapid and structured approach for evaluating gene-based testing. *Genet Med*. 2007;7:473–8.
15. Burke W, Zimmern R. Moving beyond ACCE: an expanded framework for genetic test evaluation. Cambridge, UK: PHG Foundation; 2007.
16. Aretz S, Rautenstraub B, Timmerman V. Indication criteria for genetic testing. Evaluation of validity and clinical utility. Munich, Germany: German Society of Human Genetics; 2008.
17. Schmidtke J, Cassiman JJ. The EuroGentest Clinical Utility Gene Cards. *Eur J Hum Genet*. 2010;24:1. <https://doi.org/10.1038/ejhg.2010.85>.
18. Dierking A, Schmidtke J, Matthijs G, Cassiman JJ. The EuroGentest Clinical Utility Gene Cards continued. *Eur J Hum Genet*. 2013;21:1. <https://doi.org/10.1038/ejhg.2012.161>.
19. Dierking A, Schmidtke J. The future of Clinical Utility Gene Cards in the context of next-generation sequencing diagnostic panels. *Eur J Hum Genet*. 2014;22:1247. <https://doi.org/10.1038/ejhg.2014.23>.
20. Wright CF, Kroese M. Evaluation of genetic tests for susceptibility to common complex diseases: why, when and how? *Hum Genet*. 2010;127:125–34.
21. Calonge N, Green NS, Rinaldo P, et al. Committee report: method for evaluating conditions nominated for population-based screening of newborns and children. *Genet Med*. 2010;3:153–9.
22. Sun F, Bruening W, Erinoff E, Schoelles KM, ECRI Institute Evidence-based Practice Center. Addressing challenges in genetic test evaluation. Evaluation frameworks and assessment of analytic validity. Rockville, USA: Agency for Healthcare Research and Quality; 2011.
23. Canestaro WJ, Pritchard DE, Garrison LP, Dubois R, Veenstra DL. Improving the efficiency and quality of the value assessment process for companion diagnostic tests: The Companion test Assessment Tool (CAT). *J Manag Care Spec Pharm*. 2015;21:700–12.
24. Medical Service Advisory Committee. Australian Government, Department of Health. Clinical Utility Card for heritable mutations which increase risk in [disease area]. 2016. Available at [http://www.msac.gov.au/internet/msac/publishing.nsf/Content/9C7DFC1C2DD56CBECA25801000123C32/\\$File/CUC-proforma-assessment-genetic-testing.pdf](http://www.msac.gov.au/internet/msac/publishing.nsf/Content/9C7DFC1C2DD56CBECA25801000123C32/$File/CUC-proforma-assessment-genetic-testing.pdf). Accessed 18 Apr 2017.
25. Advisory Committee on New Predictive Genetic Technologies. Genetic Services in Ontario: Mapping the Future. Report of the

- Provincial Advisory Committee on New Predictive Genetic Technologies. Ontario, Canada: Advisory Committee on New Predictive Genetic Technologies; 2001.
26. Blanquaert I, Bouchard L, Chikhaoui Y, Cleret, de Langavant G. Molecular genetics viewed from the Health Technology Assessment Perspective. *Eur J Hum Genet.* 2001;1:309–10. (abstract)
 27. Blanquaert I. Testing for BRCA: the Canadian experience. In: Kroese M, Elles R, Zimmern RL, editors. *The Evaluation of Clinical Validity and Clinical Utility of Genetic Tests, Summary of an expert workshop*, 26 and 27 June 2006. Cambridge, UK: PHG Foundation; 2007. p. 23–25.
 28. Giacomini M, Miller F, Browman G. Confronting the “gray zones” of technology assessment: evaluating genetic testing services for public insurance coverage in Canada. *Int J Technol Assess Health Care.* 2003;2:301–16.
 29. Gutiérrez-Ibarluzea I Personalised Health Care, the need for reassessment. A HTA perspective far beyond cost-effectiveness. *Ital J Public Health.* 2012; <https://doi.org/10.2427/8653>.
 30. Merlin T, Farah C, Schubert C, Mitchell A, Hiller JE, Ryan P. Assessing personalized medicines in Australia: a national framework for reviewing codependent technologies. *Med Decis Mak.* 2013;3:333–42.
 31. Fleeman N, Martin Saborido C, Payne K, et al. The clinical effectiveness and cost-effectiveness of genotyping for CYP2D6 for the management of women with breast cancer treated with tamoxifen: a systematic review. *Health Technol Assess.* 2011;33:1–102.
 32. Allingham-Hawkins DJ, Lea A, Spock L, Levine S Hayes Genetic Test Evaluation (GTE) Program: Evidence-based evaluation of genetic tests. Inaugural Meeting of the Genomic Applications in Practice and Prevention Network. Ann Arbor, USA, 29-30 Oct 2009 (abstract).
 33. Betti S, Boccia A, Boccia S, et al. HTA of genetic testing for susceptibility to venous thromboembolism in Italy. *Ital J Public Health.* 2012; <https://doi.org/10.2427/6348>.
 34. Goel V. Appraising organised screening programmes for testing for genetic susceptibility to cancer. *BMJ.* 2001;7295:1174–8.
 35. Burke W, Coughlin SS, Lee NC, Weed DL, Khoury MJ. Application of population screening principles to genetic screening for adult-onset conditions. *Genet Test.* 2001;3:201–11.
 36. Andermann A, Blanquaert I, Beauchamp S, Costea I. Guiding policy decisions for genetic screening: developing a systematic and transparent approach. *Public Health Genom.* 2011;1:9–16.
 37. Andermann A, Blanquaert I, Déry V. Genetic screening: a conceptual framework for programs and policy-making. *J Health Serv Res Policy.* 2010;2:90–97.
 38. Rousseau F, Lindsay C, Charland M, et al. Development and description of GETT: a genetic testing evidence tracking tool. *Clin Chem Lab Med.* 2010;10:1397–407.
 39. Hornberger J, Doberne J, Chien R. Laboratory-developed test--SynFRAME: an approach for assessing laboratory-developed tests synthesized from prior appraisal frameworks. *Genet Test Mol Biomark.* 2012;6:605–14.
 40. Lin JS, Thompson M, Goddard KA, Piper MA, Heneghan C, Whitlock EP. Evaluating genomic tests from bench to bedside: a practical framework. *BMC Med Inform Decis Mak.* 2012;12:117.
 41. Frueh FW, Quinn B. Molecular diagnostics clinical utility strategy: a six-part framework. *Expert Rev Mol Diagn.* 2014;14:777–86.
 42. National Academies of Sciences, Engineering, and Medicine. *An evidence framework for genetic testing.* Washington, USA: The National Academies Press; 2017.
 43. Centers for Disease Control and prevention. *First ACCE Review: Population-based Prenatal Screening for Cystic Fibrosis via CarrierTesting.* 2002. Available at <https://www.cdc.gov/genomics/gtesting/acce/acce.htm>. Accessed 26 Apr 2017.
 44. UK Genetic Testing Network website. Available at <http://ukgtn.nhs.uk/find-a-test/gene-dossiers>. Accessed 8 May 2017.
 45. Evaluation of Genomic Applications in Practice and Prevention web site. Available at <http://www.cdc.gov/egappreviews/>. Accessed 8 May 2017.
 46. McClain MR, Palomaki GE, Piper M, Haddow JE. A rapid-ACCE review of CYP2C9 and VKORC1 alleles testing to inform warfarin dosing in adults at elevated risk for thrombotic events to avoid serious bleeding. *Genet Med.* 2008;2:89–98.
 47. German Society of Human Genetics website. Available at http://www.gfhev.de/de/leitlinien/Diagnostik_LL.htm. Accessed 8 May 2017.
 48. EuroGentest website. Available at <http://www.eurogentest.org/index.php?id=668>. Accessed 21 Sept 2017.
 49. European Journal of Human Genetics website. Available at http://www.nature.com/ejhg/archive/categ_genecard_012017.html?lang=en. Accessed 8 May 2017.
 50. Advisory Committee on Heritable Disorders in Newborns and Children web site. Available at <http://www.hrsa.gov/advisorycommittees/mchbadvisory/heritabledisorders>. Accessed 26 Aug 2017.
 51. Segal JB, Brotman DJ, Emadi A, et al. Outcomes of genetic testing in adults with a history of venous thromboembolism. Evidence Reports/Technology Assessments, No. 180. Rockville, USA: Agency for Healthcare Research and Quality. 2009.
 52. Medical Service advisory committee website. Available at <http://www.health.gov.au/internet/msac/publishing.nsf/Content/1411-public>. Accessed 05 Apr 2017.
 53. Tranchemontagne J, Boothroyd L, Blanquaert I. Contribution of BRCA1/2 Mutation Testing to Risk Assessment for Susceptibility to Breast and Ovarian Cancer. Monograph. Montréal, Canada: Agence d'évaluation des technologies et des modes d'intervention en santé; 2006.
 54. Hayes Inc. Genetic test evaluation website. Available at <https://www.hayesinc.com/hayes/publications/genetic-test-evaluation>. Accessed 25 Apr 2017.
 55. Zimmern RL, Kroese M. The evaluation of genetic tests. *J Public Health.* 2007;29:246–50.
 56. Grosse SD, Khoury MJ. What is the clinical utility of genetic testing? *Genet Med.* 2006;8:448–50.
 57. Kohler JN, Turbitt E, Biesecker BB. Personal utility in genomic testing: a systematic literature review. *Eur J Hum Genet.* 2017;25:662–8. <https://doi.org/10.1038/ejhg.2017.10>.
 58. Drummond MF, Sculpher MJ, Claxton K, Torrance GW, Stoddart GL, editors. *Methods for the economic evaluation of health care programmes.* 4th edn. Oxford, UK: Oxford University Press; 2015.
 59. D'Andrea E, Marzuillo C, De Vito C, et al. Which BRCA genetic testing programs are ready for implementation in health care? A systematic review of economic evaluations. *Genet Med.* 2016;18:1171–80.
 60. NHS UKGTN Testing Criteria. Available at https://ukgtn.nhs.uk/fileadmin/uploads/ukgtn/Documents/Resources/Library/Policies_Procedures/Testing_Criteria_paper.pdf. Accessed 30 Oct 2017.
 61. Gagnon MP, Desmartis M, Lepage-Savary D, et al. Introducing patients' and the public's perspectives to health technology assessment: a systematic review of international experiences. *Int J Technol Assess Health Care.* 2011;1:31–42.
 62. HIScreenDiag Report Summary. Building a tool to evaluate and improve health investments in screening and diagnosis of disease. 2011. Available at http://cordis.europa.eu/result/rcn/56825_en.html. Accessed 8 May 2017.

63. EUnetHTA Joint Action 2, Work Package 8. HTA Core Model @ version 3.0. Available at <https://www.htacoremodel.info/BrowseModel.aspx>. Accessed 8 May 2017.
64. Guyatt GH, Oxman AD, Vist GE, GRADE Working Group. et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. 2008;336:924–6.
65. UK Genetic Testing Network. Fifth report of the UKGTN. Promoting gene testing together. London, UK: UK Genetic Testing Network; 2017.
66. Matthijs G, Dierking A, Schmidtke J. New EuroGentest/ESHG guidelines and a new clinical utility gene card format for NGS-based testing. *Eur J Hum Genet*. 2016;24:1. <https://doi.org/10.1038/ejhg.2015.229>.