phg foundation
making science work for health

Testing Challenges: The Clinical Use of Genetic and Molecular Biomarkers

A Public Health Perspective

CAPABILITY *Second Workshop*

May 21 - 23, 2008

Oxford, UK

phg
making science work for health

Contents of Lecture

1. The problem
2. Conceptual issues
2. Evaluation
3. Regulation

The General Thesis

1. **Problems and issues concerning the evaluation and regulation of genetic tests are generic, applicable to all forms of diagnostics and biomarkers**

2. **The failure to address these matters will be of major public health concern in future years**

phg
making science work for health

The Context

1. **The completion of the Human Genome Project, new technology and advances in cell and molecular biology have together led to the development of new tests and biomarkers at an unprecedented rate**
2. **These tests are now more complex than ever before, both in terms of the technologies used and in their interpretation**
3. **They are being made more generally available – to non specialists and direct to the public**
4. **The assessment of predictive or susceptibility tests brings its own challenges – it is not entirely practical or feasible to wait many years before outcome is definitively known**

5 **Existing regulatory and evaluative mechanisms carried out under the European Directive on In Vitro Devices are primarily concerned with the safety of devices and assays and the assessment of analytical validity**

6. **Commissioners, funders or reimbursers of health services are all under extreme financial pressure and require evidence of effectiveness before they will consider investment in the test**

Conceptual Issues

Diagnosis

What is diagnosis?

The crucial process that labels patients and classifies their illnesses, that identifies (and sometimes seals) their likely fates or prognoses, and that propels us toward specific treatments in the confidence (often unfounded) that they will do more good than harm.

David Sackett (1991) Clinical Epidemiology: A Basic Science for Clinical Medicine

The label - the diagnosis - is not an end in itself but an intermediary, a means to an end

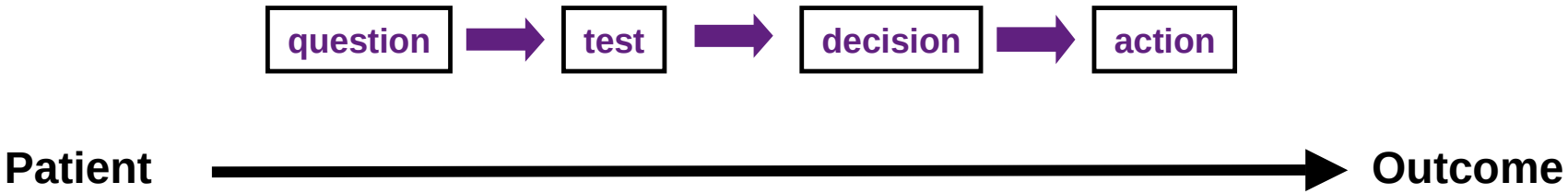
phg
making science work for health

Purpose or Uses of a Test or Biomarker

1. Diagnosis
2. Risk stratification
3. Disease prognosis
4. Treatment stratification
5. Treatment monitoring
6. Population screening

Why Do A Test?

Purpose is all important



After Price CP & Christenson RH (2007) The Clinical Question: A System for Formulating Answerable Questions in Laboratory Medicine. In Evidence Based Laboratory Medicine. Ed: Price CP and Christenson RH.

phg
making science work for health

Effectiveness

The effectiveness of an intervention is the extent to which it achieves the objective (purpose) for which it was designed

phg
making science work for health

Definition of Biomarker

A characteristic that is objectively measured and evaluated as an indicator of normal biologic processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention

Types of Tests or Biomarkers

1. Clinical
2. Imaging
3. Genetic
4. Cellular and molecular
5. Physiological
6. Psycho-social

Assays and Tests

Assay

A method for determining the presence or quantity of a component

Test

A procedure that makes use of an assay for a particular purpose

Tests -The Importance of Context

CONTEXT MATTERS IN DECIDING THE EFFECTIVENESS OF A TEST

The term test is used as a shorthand for referring to an assay used in the context of:

1. **a particular disease**
2. **in a particular population**
3. **for a particular purpose**

An allternative conceptualisation is to treat the *assay* as the measurement and the *test* as the interpretation of that measurement

phg
making science work for health

Implications of the Assay-Test Distinction

The practical implication of the distinction is that whereas the evaluation of an assay is reasonably straightforward and allows broadly applicable standards to be established, the evaluation of a test is more complex and inherently less susceptible to standardisation.

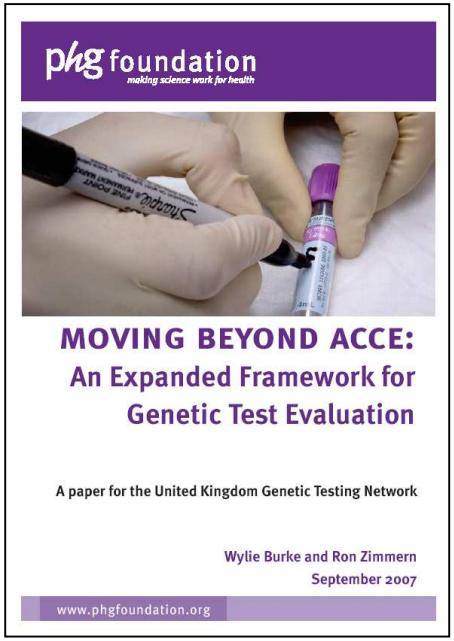
Each test is likely to need evaluation in its individual context, depending on disease, purpose and population

phg
making science work for health

Evaluation

phg
making science
work for health

The ACCE Framework



1. **A** nalytical validity
2. **C** linical validity
3. **C** linical utility
4. **E** thical, legal and social

Analytical validity of a test defines its ability to measure accurately and reliably the **component of interest**

Clinical validity of a test defines its ability to detect or predict the presence or absence of clinical disease or predisposition to disease

Clinical utility of a test refers to the likelihood that the test will lead to an improved outcome

Ethical, legal and social implications of a test

The ACCE framework is applicable to all forms of molecular diagnostics and biomarkers

phg
making science work for health

Framework for Test Evaluation (1)

Domain	Specific Element	Focus of evaluation
Pre-evaluation definition	***Test Definition***	Precise definition of: Genetic variants to be assayed Disorder Population Purpose
Assay		
	Analytic validity	Sensitivity Specificity PPV, NPV
	Reliability and Reproducibility	Kappa
Clinical Validity		
	Gene-Disease Association	Primary research Systematic review Meta-analysis
	Clinical Test Performance	Sensitivity, Specificity PPV, NPV, LR+, LR-, ROC

phg
making science work for health

Clinical Validity

Scientific validity

Evaluation of the relationship between biomarker and disease

Clinical Performance

Evaluation of the test performance in the clinical situation

Evidence of biomarker-disease association is necessary, but by no means sufficient, as an indicator of effective and useful clinical performance

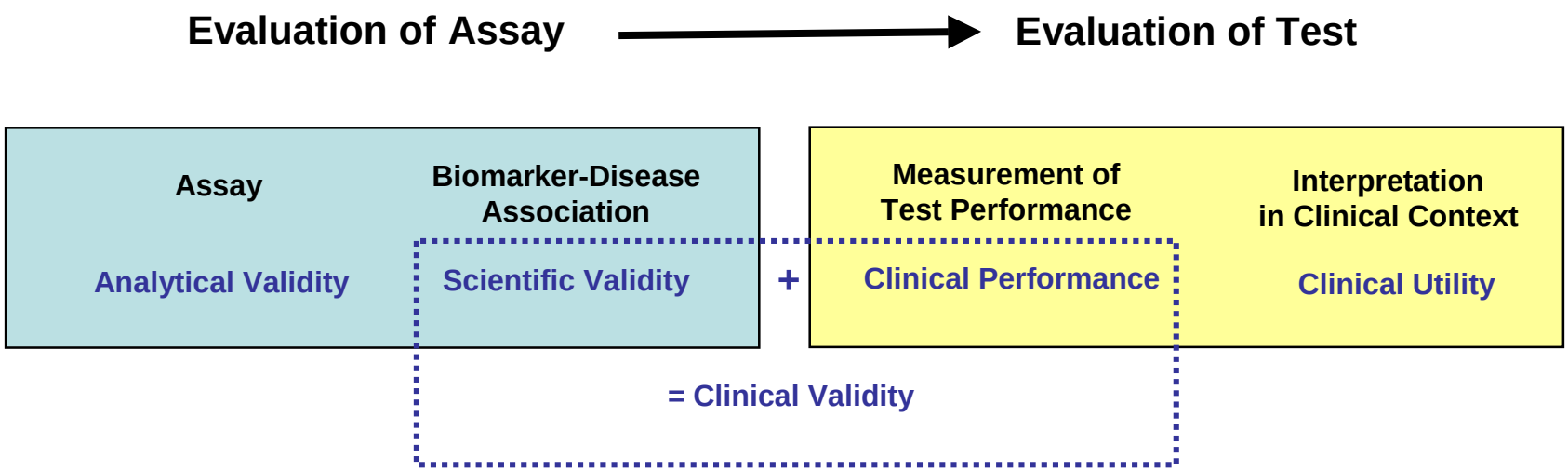
phg
making science work for health

Dimensions of Clinical Utility

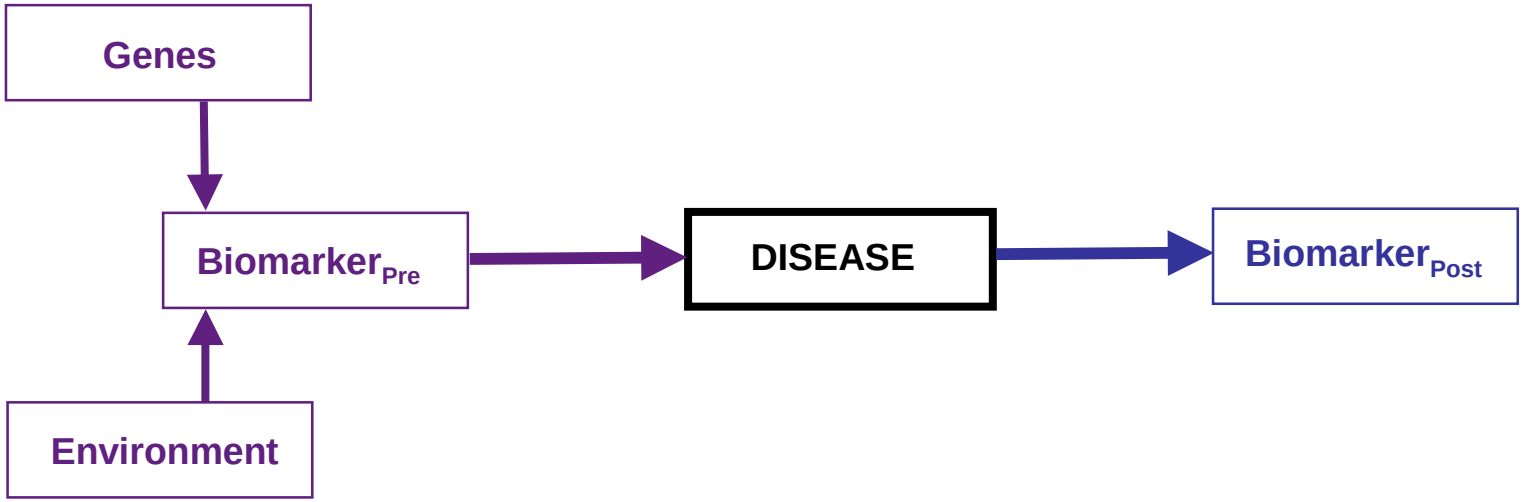
Clinical Utility		
Test Purpose	***Legitimacy***	Conformity to the social preferences expressed in ethical principles, values, norms, mores, laws and regulations
	Efficacy	Potential of test and associated services to deliver health benefit
	Effectiveness	Actual delivery of health benefit in routine clinical setting
	Appropriateness	Expected health benefit exceeds expected negative cons quences by a sufficiently wide margin that the test is worth doing
Feasibility of Test Delivery	***Acceptability***	Conformity to the wishes, desires, and expectations of patients and their families
	Economic ***Efficiency***	Ability to lower the costs of care without diminishing benefits
	Optimality	Balancing improvements in health against costs of improvements
	Equity	Just and fair distribution of health care and its benefits among members of the population.

phg
making science work for health

An Alternative Conceptualisation



Diagnostic and Predictive Tests

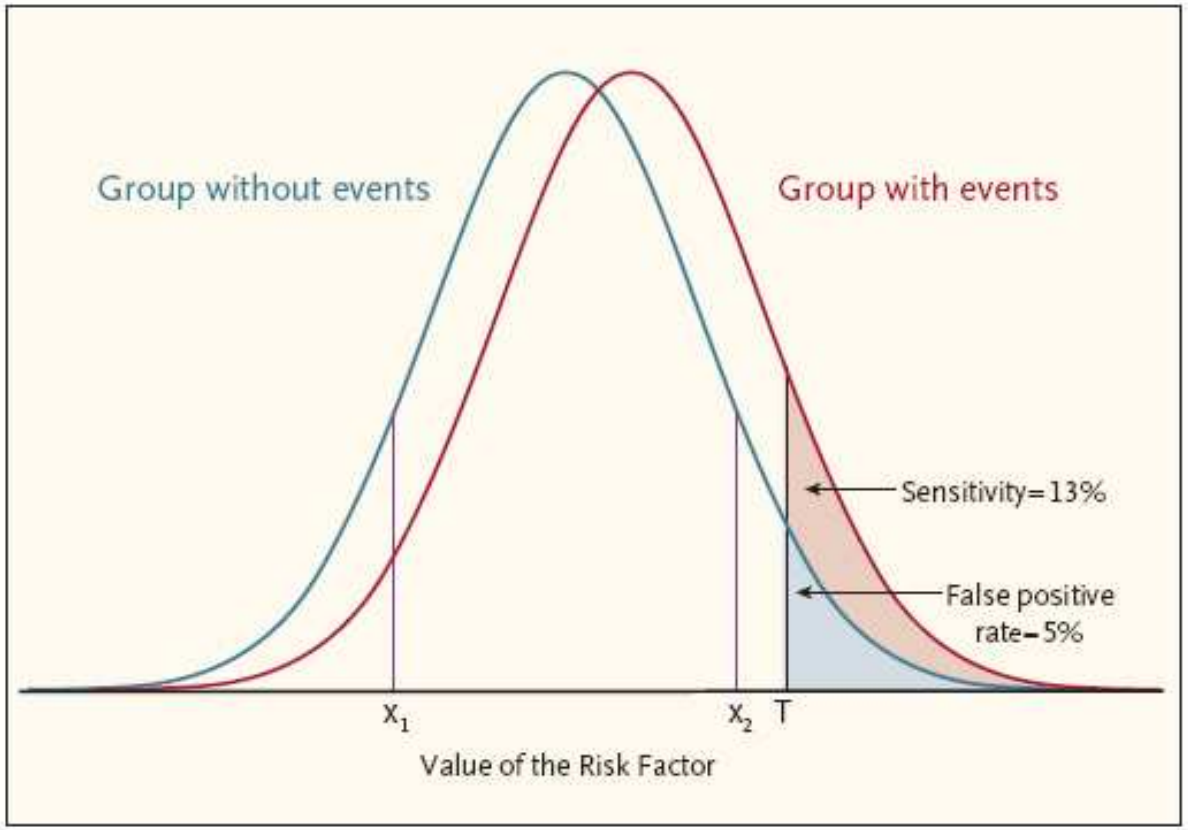


Predict future risk of disease
Monitor risk
Intervene to prevent disease

Diagnose disease
Follow course of disease
Monitor treatment

phg
making science work for health

Standard Conceptualisation of Predictive Tests

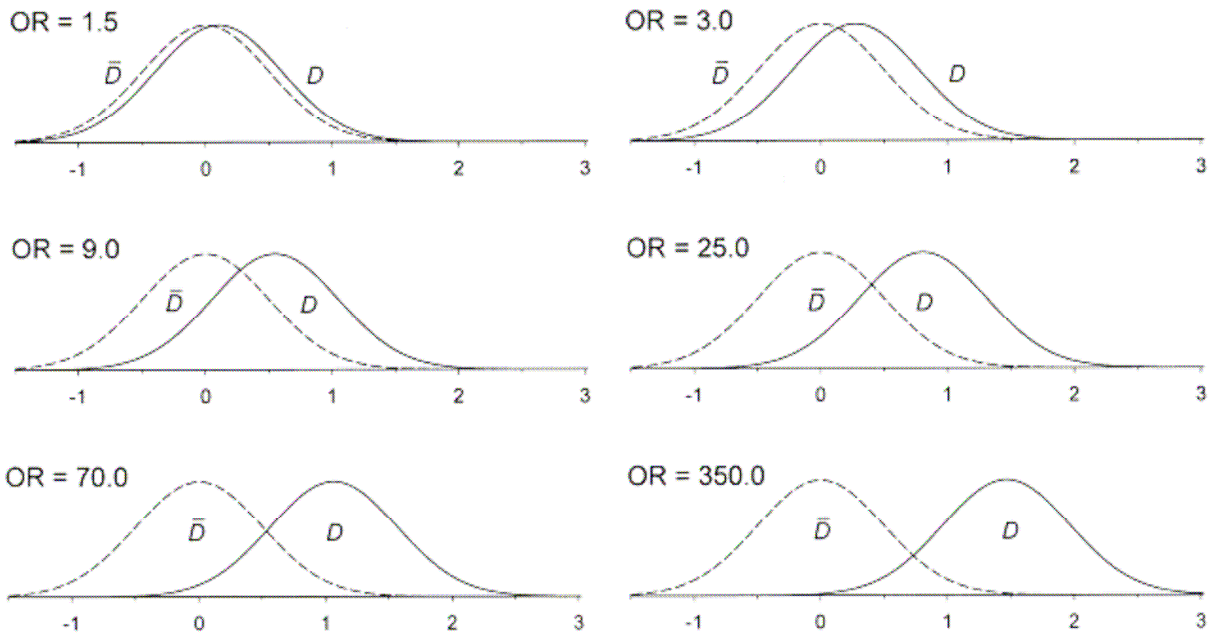


Normal Probability Density Functions of the Risk Factor among Persons Who Will Not Have the Event (Blue) and among Those Who Will (Red).

Ware (2006) NEJM 355, 2615-17

phg
making science work for health

Odds Ratios and Discrimination



Pepe et al (2004) Am J Epidemiol 159, 882-890

phg
making science work for health

Cumulative Effect for Risk of Prostate Cancer – Genotypes and Family History

Variable	Case Subjects	Control Subjects	Regression Coefficient	Odds Ratio (95% CI)	P Value†	P Value for Trend‡
Age			0.01	1.01 (1.00–1.02)	0.02	
Geographic region			–0.75	0.47 (0.40–0.55)	<0.001	
No. of associated factors**						
0	144 (5.0)	174 (10.1)	NA	1.00		
1	778 (26.9)	581 (33.6)	0.48	1.62 (1.27–2.08)	1.27×10^{-4}	
2	1053 (36.4)	622 (36.0)	0.73	2.07 (1.62–2.64)	5.86×10^{-9}	
3	642 (22.2)	286 (16.6)	0.99	2.71 (2.08–3.53)	9.54×10^{-14}	
4	236 (8.2)	60 (3.5)	1.56	4.76 (3.31–6.84)	9.17×10^{-19}	
≥5	40 (1.4)	5 (0.3)	2.24	9.46 (3.62–24.72)	1.29×10^{-8}	4.78×10^{-28}

From Zheng et al (2008) NEJM Feb 13

phg
making science work for health

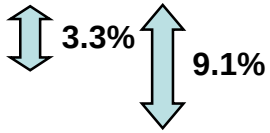
TCF7L2 and Type II Diabetes

SNP rs 7903146

Allele RR	RR (Heterozygous)	RR (Homozygous)
1.36 (1.24-1.48)	1.35 (1.19-1.43)	1.90 (1.54-2.33)

Groves et al (2006) Diabetes 55, 2640-2644

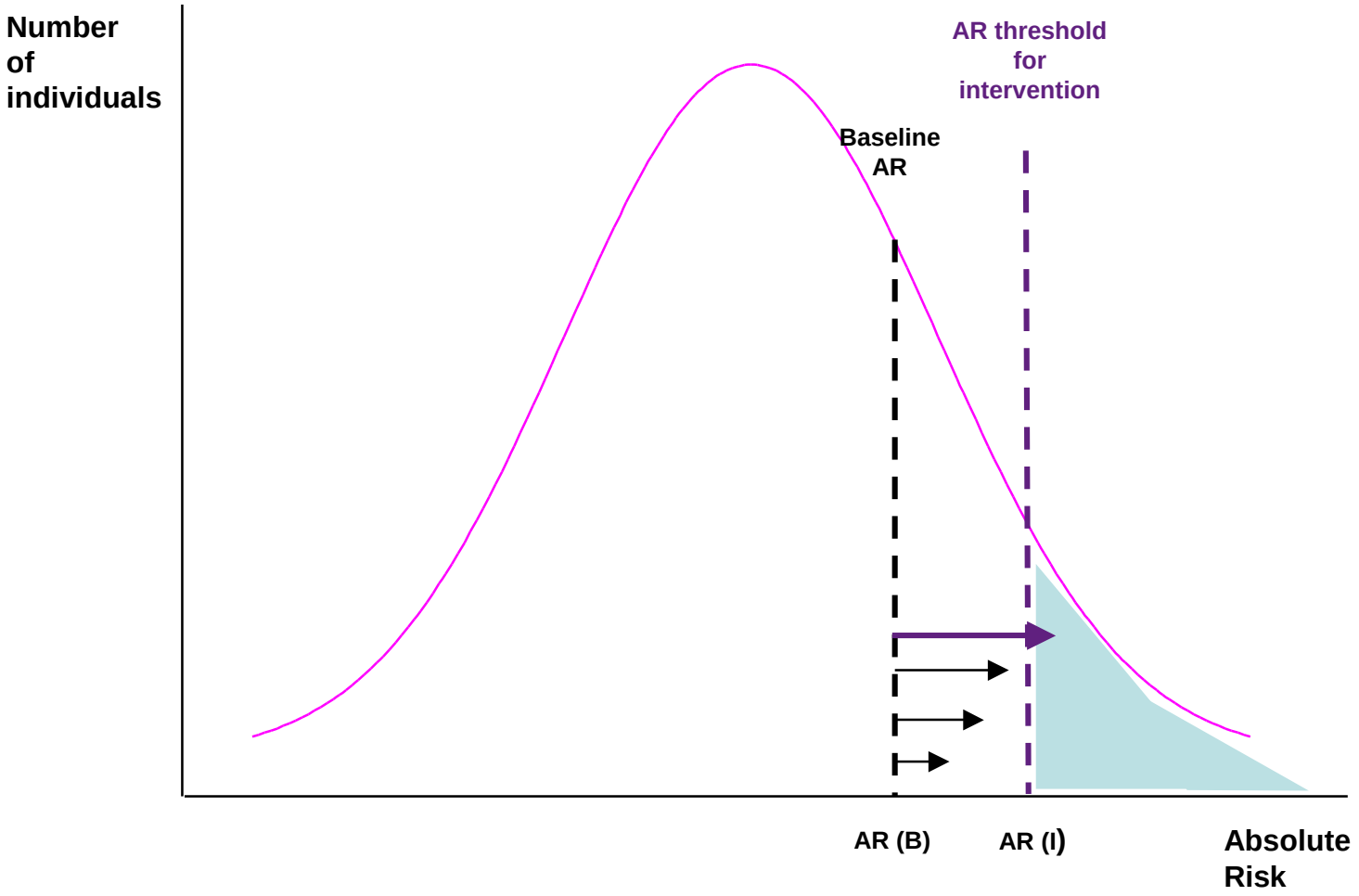
Genotype	Proportion With diabetes or IFG	Odds Ratio
Wildtype	23.34%	1.00 (0.93-1.76)
Heterozygous	26.65%	1.28 (1.05-2.65)
Homozygous	32.43%	1.67 (1.04-1.60)



Melzer et al (2006) BMC Medicine 4

phg
making science work for health

Alternative Conceptualisation – The Use of Absolute Risk



phg
making science work for health

Use of Absolute Risk

1. **Baseline absolute risk**
2. **(a) Addition to existing algorithm or**

   (b) Substitute for existing algorithm

   for determination of absolute risk
3. **Absolute risk thresehold for appropriatenes for a particular intervention**

phg
making science work for health

Regulation

phg
making science
work for health

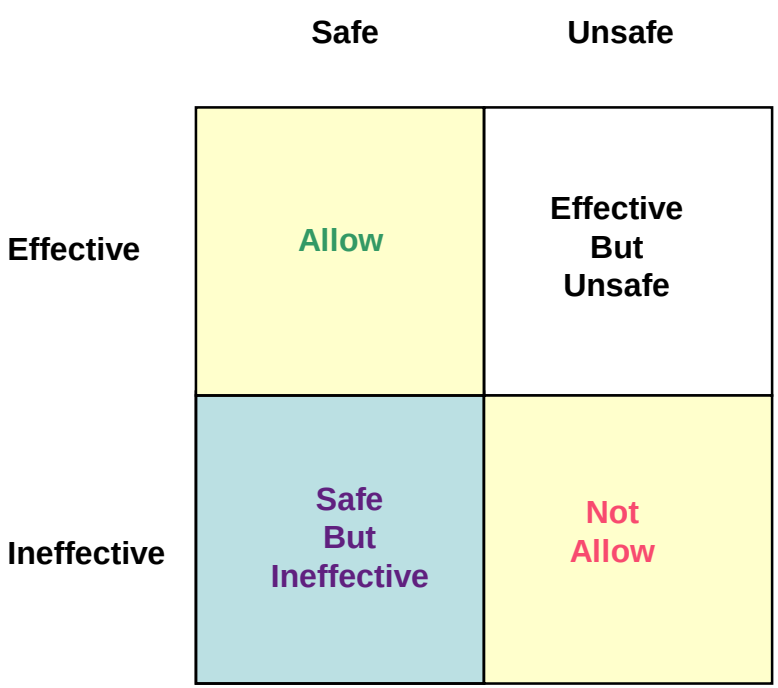
Levels of Regulation

1. **Statutory**
   - legislation
   - regulation
   - codes of practice
2. **Resources**
   - insurers
   - commissioners
   - health maintenance organisations
3. **Clinical**
   - clinical governance
   - physician and patient education

After Burke & Zimmern (2004) Nature Reviews Genetics 5, 955

phg
making science work for health

The Fundamental Issues for Regulators



How much of test regulation should be carried out in the statutory sector and how much through 'informal' mechanisms?

Should statutory regulators concern themselves with matters of clinical valdity or utility?

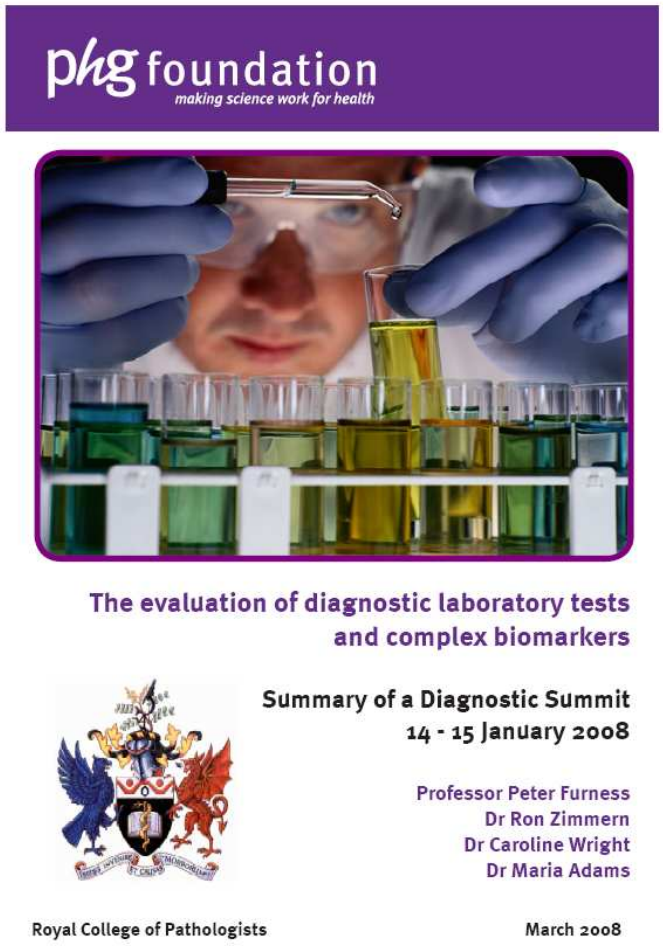
Should statutory regulators concern themselves with tests that are safe but ineffective?

How should the idea of 'safety' be interpreted in the context of tests?

Is it possible to establish a risk based system of statutory regulation?

phg
making science work for health

Diagnostic Summit



1. **A new body should be established to ensure the evaluation of diagnostic tests.**
2. **A publically available database be created of new and existing laboratory tests – a 'diagnostics formulary' – containing evidence for clinical performance, and explicitly stating where any evidence is lacking.**
3. **Policy makers and industry should be encouraged to address issues around gathering the necessary evidence for clinical evaluation.**
4. **An independent expert body should be responsible for evaluating the evidence for test performance and for making recommendations about appropriate clinical use.**
5. **Commissioners and health care professionals should be encouraged to use only those tests where appropriate evidence of clinical performance exists.**
6. **Statutory regulators should be empowered to require transparency relating to evidence of test performance, and ensure responsive and proportionate risk assessment to ensure patient safety.**

phg
making science work for health

Direct to Consumer Company Survey

Internet search for direct-to-consumer (DTC) tests on human nucleic acids (DNA or RNA) for disease using a sample obtained non-invasively

29 companies found (80% in US)

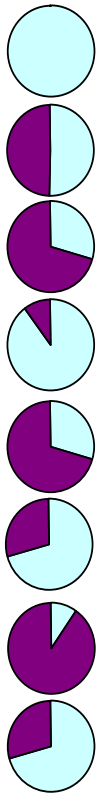
23andMe, Acu-Gen Biolab, Consumer Genetics, CygeneDirect, deCODE genetics, Dermagenetics, DNAdirect, Genelex, Genetic health, Genosolutions, G-nostics, Graceful Earth, Health Tests Direct, HealthCheckUSA, Holistic Heal, Integrative Genomics, Interleukin Genetics, Kimball Genetics, Knome, Market America, Medi-Checks, Medigenomix, Mygenome.com, Navigenics, Quitxar, Salugen, Sciona, Smart Genetics, Suracell

Type of test: single PCR tests, specific gene sequencing, microarrays, DNA methylation, genome-wide scans, complete genome sequencing

Purpose of test: diagnosis (e.g. carrier screening), susceptibility to complex diseases, nutrigenomics, pharmacogenetics, lifestyle factors

Assessed against American Society of Human Genetics (ASHG) recommendations for DTC genetic tests*

ASHG Recommendation	Approximate proportion of companies meeting recommendation
1. Clinical validity	0%
2, i. Scientific publications	50%
ii. Details of test provided	70%
3. Risks associated	10%
4. i. Laboratory certification	70%
ii. Analytical validity	30%
5. Privacy policy	90%
6. Evidence of intervention (where recommended)	30%



phg
making science work for health

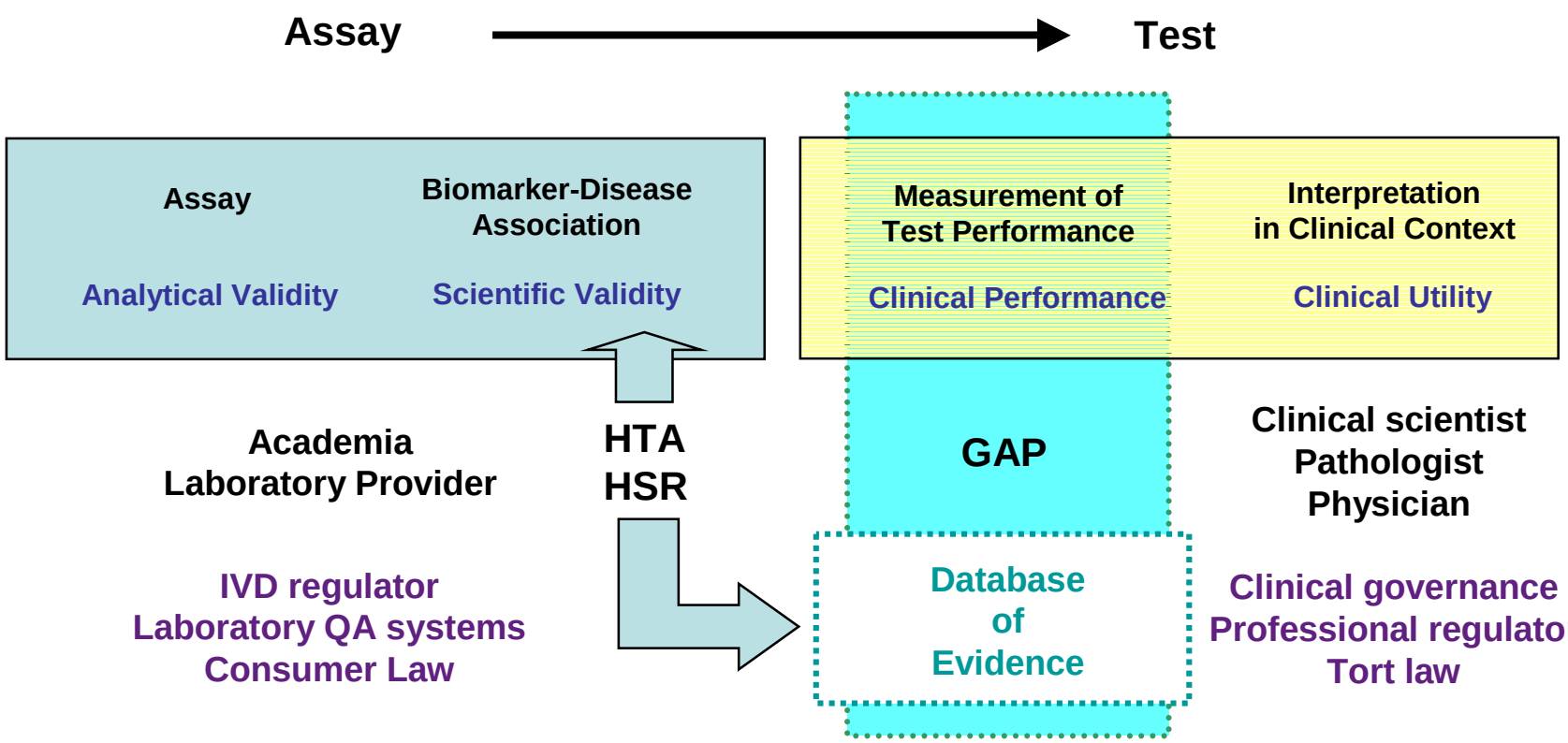
Epidemiological Study of DTC Tests

1. Seven companies offering predictive testing using multiple markers involving 69 polymorphisms in 56 genes
2. Literature review on 260 meta-analyses addressed 46 of the 69 polymorphisms and 32 of the 56 genes, encompassing 160 unique polymorphism-disease asssociations
3. Statistically significant associations were only found in 60 (38%) of these 160. These involved 29 polymorphisms and 28 different diseases
4. The odds ratios ranged from 0.54 to 0.88 for protective associations and from 1,04 to 3.2 for risk variants
5. The main commonly studied polymorphisms were found in the genes MTHFR, TNF-alpha, GSTP1, GSTT1 and VDR

Cecile Janssens et al (2008) Am J Hum Gen. 82, 593-599

phg
making science work for health

Regulation of Tests and Biomarkers



phg
making science work for health

Conclusion

1. **The distinction between an ‘assay’ and a ‘test’ is crucial to understanding the roles of statutory regulators and others in test evaluation**
2. **Clinical validity requires more than evidence of gene-disease association; it also requires evidence of test performance such as sensitivity, specificity, positive and negative predictive values**
3. **Main problem is lack of data; policy is urgently needed to establish systems and resources to generate evidence of test performance, and to agree the respective roles and responsibilities of government, statutory regulators, public bodies, academia and the commercial sector**
4. **Systems should be established to ensure that the data are appropriately anaysed and evaluated against agreed standards and that the evidence is placed in the public domain**
5. **Funders and reimbursers of health services and clinicians should be discouraged from using tests that are not backed by appropriate clinical evidence**
6. **The role of statutory regulators should be confined to ensuring the safety of all tests and biomarkers, and that evidence in relation to test performance (or lack of it) is placed in the public domain**