

Öffentliche Anhörung des Deutschen Ethikrates  
Berlin, 3. Mai 2012

# Möglichkeiten und Grenzen prädiktiver genetischer Diagnostik multifaktorieller Erkrankungen

Thomas F. Wienker ([wienker@molgen.mpg.de](mailto:wienker@molgen.mpg.de))  
Max-Planck-Institut für Molekulare Genetik, Berlin-Dahlem



## **Krankheitsprognose durch DNA-Lesen begrenzt**

**Baltimore/Washington** Aus dem gesamten Erbgut eines Menschen lässt sich nur begrenzt das Risiko für häufige Krankheiten wie Krebs vorhersagen. Zu diesem Ergebnis kommen US-Forscher nach der Auswertung Tausender Daten von eineiigen Zwillingen aus früheren Erhebungen.

«Wir glauben, dass Genomtests kein Ersatz für gängige Strategien zur Krankheitsvorbeugung sein werden», sagte Bert Vogelstein vom Johns Hopkins Kimmel Cancer Center (Baltimore, USA). Das Team um Vogelstein, Kenneth Kinzler und Nicholas J. Roberts präsentiert seine Daten im Fachjournal «Science Translational Medicine» und auf einer Krebstagung in Chicago.

**A complete electronic version of this article** and other services, including high-resolution figures, can be found at:

<http://stm.sciencemag.org/content/early/2012/04/25/scitranslmed.3003380.1>



SCIENTIST AT WORK | DAVID B. GOLDSTEIN

# A Dissenting Voice as the Genome Is Sifted to Fight Disease

By NICHOLAS WADE  
Published: September 15, 2008

---

**The New York Times**

The common disease/common variant idea is largely wrong. What has happened is that a multitude of rare variants lie at the root of most common diseases, being rigorously pruned away as soon as any starts to become widespread.

“There is absolutely no question,” he said, “that for the whole hope of personalized medicine, the news has been just about as bleak as it could be.”

---

## Common SNPs explain a large proportion of the heritability for human height

Jian Yang<sup>1</sup>, Beben Benyamin<sup>1</sup>, Brian P McEvoy<sup>1</sup>, Scott Gordon<sup>1</sup>, Anjali K Henders<sup>1</sup>, Dale R Nyholt<sup>1</sup>, Pamela A Madden<sup>2</sup>, Andrew C Heath<sup>2</sup>, Nicholas G Martin<sup>1</sup>, Grant W Montgomery<sup>1</sup>, Michael E Goddard<sup>3</sup> & Peter M Visscher<sup>1</sup>

**... We show that 45% of variance can be explained by considering all SNPs simultaneously. Thus, most of the heritability is not missing but has not previously been detected because the individual effects are too small to pass stringent significance tests. ...**

---

# A Commentary on ‘Common SNPs Explain a Large Proportion of the Heritability for Human Height’ by Yang et al. (2010)

---

Peter M. Visscher,<sup>1</sup> Jian Yang<sup>1</sup> and Michael E. Goddard<sup>2,3</sup>

<sup>1</sup> Queensland Statistical Genetics Laboratory, Queensland Institute of Medical Research, Brisbane, Australia

<sup>2</sup> Department of Food and Agricultural Systems, University of Melbourne, Australia

<sup>3</sup> Biosciences Research Division, Department of Primary Industries, Victoria, Melbourne Australia

... following the publication of Yang et al. (2010) it became clear to us that the methodology we applied, the interpretation of the results and the consequences of the findings on the genetic architecture of human height and that for other traits such as complex disease are not well understood or appreciated. Here we explain some of these issues ...

## Hundreds of variants clustered in genomic loci and biological pathways affect human height

A full list of authors and their affiliations appears at the end of the paper.

... Here, using 183,727 individuals, we show that hundreds of genetic variants, in at least 180 loci, influence adult height, a highly heritable and classic polygenic trait<sup>2,3</sup>. The large number of loci reveals patterns with important implications for genetic studies of common human diseases and traits. ...

# Genetic Heterogeneity in Human Disease

Jon McClellan<sup>1,\*</sup> and Mary-Claire King<sup>2,\*</sup>

<sup>1</sup>Department of Psychiatry

<sup>2</sup>Departments of Medicine and Genome Sciences

University of Washington, Seattle, WA 98195-7720, USA

\*Correspondence: drjack@uw.edu (J.M.), mcking@uw.edu (M.-C.K.)

DOI 10.1016/j.cell.2010.03.032

Strong evidence suggests that rare mutations of severe effect are responsible for a substantial portion of complex human disease. Evolutionary forces generate vast genetic heterogeneity in human illness by introducing many new variants in each generation. Current sequencing technologies offer the possibility of finding rare disease-causing mutations and the genes that harbor them.

“Every unhappy family is unhappy in its own way,” wrote Tolstoy in *Anna Karenina*. Tolstoy was reflecting on the individually unique nature of human tragedy. We suggest that this principle also captures the misfortune of human disease.

# Genetische Heterogenität: Locus Heterogenität & allelische Heterogenität

- **Locus Heterogenität**

veraltet: (echte) Heterogenie

Varianten an unterschiedlichen Genorten sind ursächlich für eine klinisch  $\pm$  einheitliche Krankheit;

- **allelische Heterogenität**

erzeugt eine „allelische Serie“

unterschiedliche Varianten ( $\sim$ Mutationen) an ein und demselben Genort sind ursächlich für eine klinisch  $\pm$  einheitliche Krankheit;

> ***intrafamiliär einförmig, interfamiliär u. U. variabel***



## **‘Phantom Heritability’ Indicates Poor Predictive Value of Gene Tests**

Helen Wallace, January 10, 2012

[This] paper undermines frequently reported claims that sequencing the human genome will lead to a “genetic revolution” in healthcare, in which the entire population will have their genomes mapped or sequenced in order to predict and prevent the diseases they will get. It casts doubt on this idea because it recognizes that **common diseases are in fact more complex** than enthusiasts for this approach have previously admitted: meaning that there will be **inherent, biological limitations to the predictive value** of whole genome sequencing.

# The mystery of missing heritability: Genetic interactions create phantom heritability

Or Zuk<sup>a</sup>, Eliana Hechter<sup>a</sup>, Shamil R. Sunyaev<sup>a,b</sup>, and Eric S. Lander<sup>a,1</sup>

Broad Institute of MIT and Harvard, Cambridge, MA 02142; and <sup>b</sup>Genetics Division, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115

Contributed by Eric S. Lander, December 5, 2011 (sent for review October 9, 2011)

Human genetics has been haunted by the mystery of “missing heritability” of common traits. Although studies have discovered >1,200 variants associated with common diseases and traits, these variants typically appear to explain only a minority of the heritability.

... Ultimately, the most important goal for biomedical research is not explaining heritability—that is, predicting personalized patient risk—but understanding pathways underlying disease and using that knowledge to develop strategies for therapy and prevention.

# GWAS

genome-wide associations studies

# Was ist signifikant ?

- statistische Signifikanz
- biologische/klinische Relevanz



Statistische Signifikanz bedeutet:

Es ist unwahrscheinlich, dass die Befunde  
durch Zufallseffekte bedingt sind

## Der Statistische Test

- Daten, die an einer Stichprobe des Umfanges  $N$  erhoben werden
- Signifikanzschwelle  $\alpha$ , üblicherweise 5%, d.h. 0,05; muss ggf. bei mehrfacher Testdurchführung adjustiert werden (multiples Testen)
- Null-Hypothese,  $H_0$ ; üblicherweise:  
die beobachteten Daten sind nur durch Zufallseffekte bedingt
- P-Wert: Wahrscheinlichkeit, dass die Daten so sind wie sie sind, wenn die Null-Hypothese gilt,  $\text{Prob}(\text{Data} \mid H_0 = \text{true})$
- Signifikanz wird erklärt, wenn  $P \leq \alpha$

## Kleine P-Werte bedeuten statistische Signifikanz

Der P-Wert ist abhängig von  
Effektstärke und Stichprobenumfang

- grosse Effekte bedingen kleine P-Werte;
- kleine Effekte können bei grossen Stichproben signifikante P-Werte ergeben

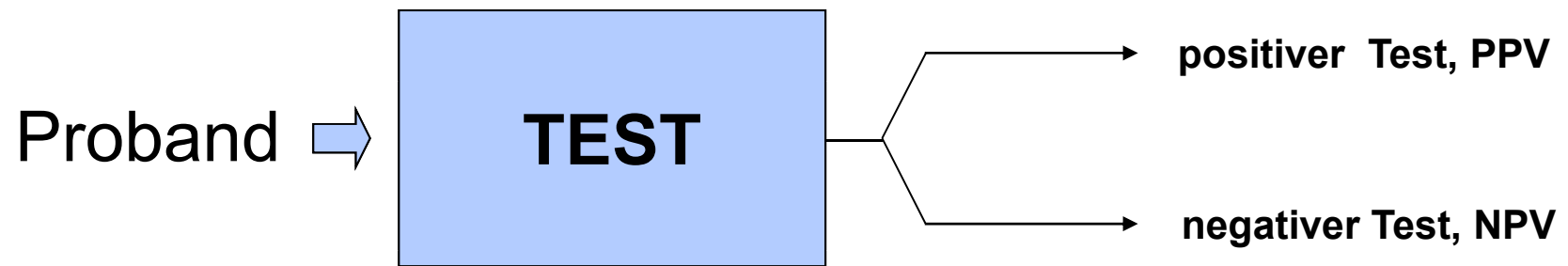
## GWAS (genome-wide association studies)

Sehr grosse Stichproben (2.000 bis über 100.000) bei kleinen Effekten, welche biologisch/klinisch bedeutungslos sind, ergeben kleine P-Werte, also hohe (statistische) Signifikanz.

Bei etablierter statistischer Signifikanz  
sind P-Werte **irrelevant**  
für die Beurteilung der biologischen/klinischen Relevanz.

# Prädiktiver / diagnostischer Test

---



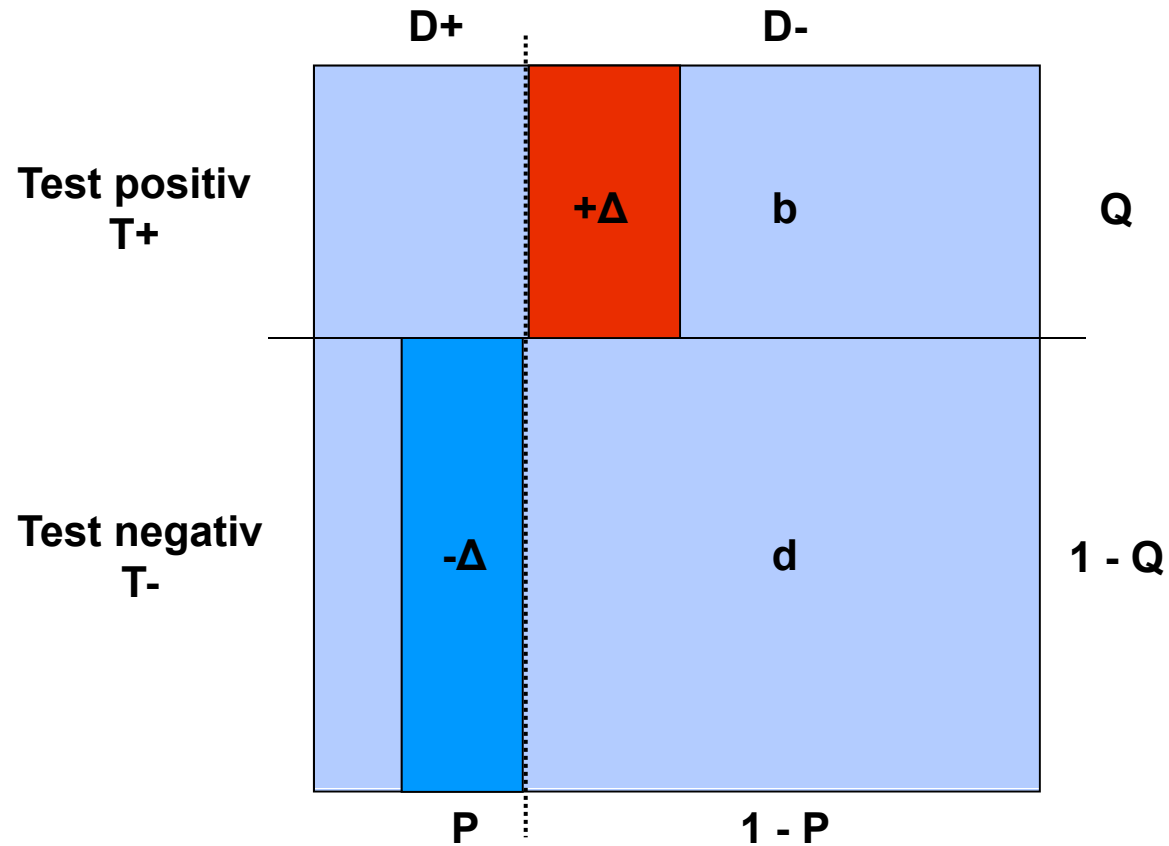
# prädiktiver / diagnostischer Test

	<b>D+</b>	<b>D-</b>		<b>D : Disease +/-</b>
<b>Test positiv T+</b>	<b>a</b>	<b>b</b>	<b>Q</b>	<b>Q : Frequenz des positiven Tests</b>
<b>Test negativ T-</b>	<b>c</b>	<b>d</b>	<b>1 - Q</b>	
	<b>P</b>	<b>1 - P</b>		<b>P : Prävalenz</b>

$$a = P \cdot Q \quad , \quad b = (1 - P) \cdot Q \quad , \quad c = P \cdot (1 - Q) \quad , \quad d = (1 - P) \cdot (1 - Q)$$



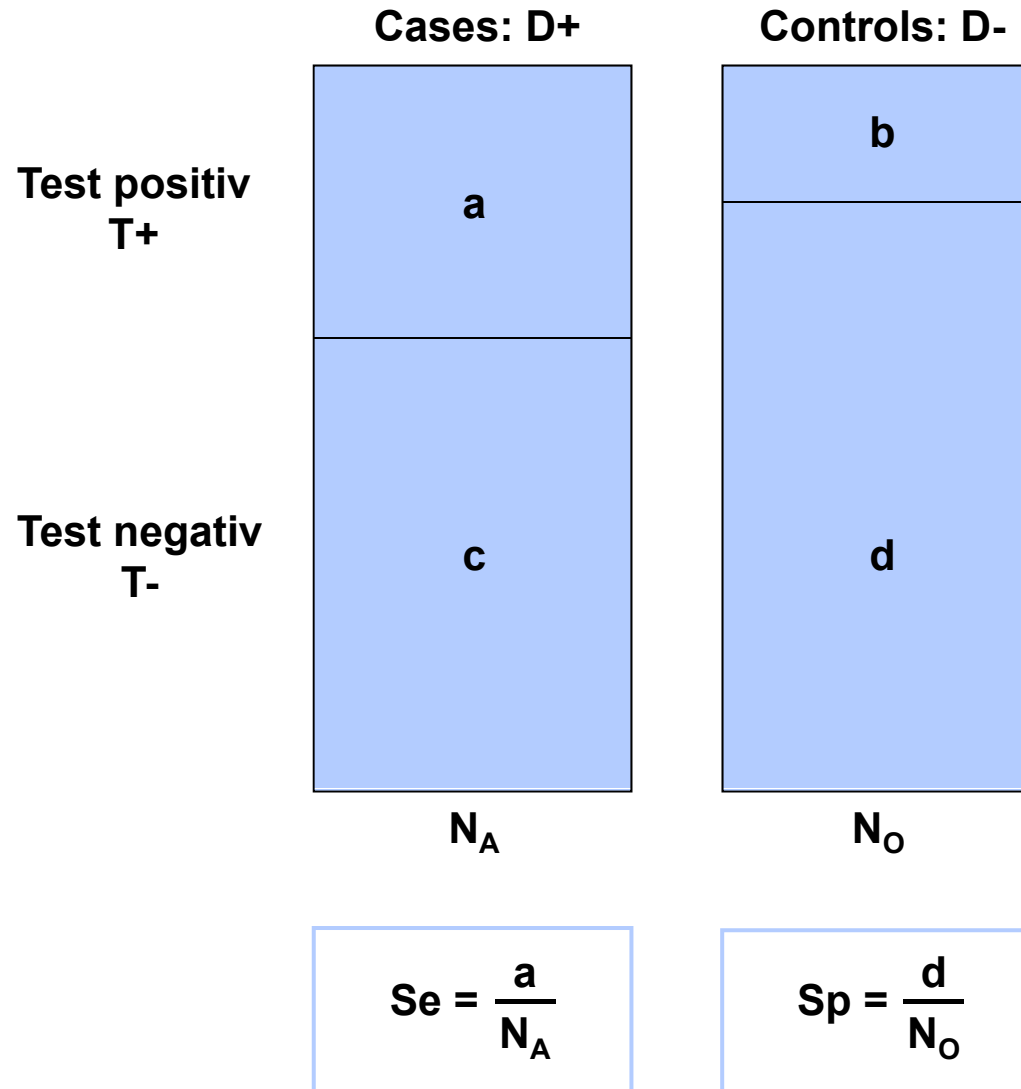
# prädiktiver / diagnostischer Test



$\Delta$  : redistributed area

# prädiktiver / diagnostischer Test

---



		Status: krank / gesund		Summe
		D+ „Kranke“	D- „Gesunde“	
		Prävalenz: P	1 - P = P'	1.0
Test Ausgang	„positiv“ <b>T+</b> Test level: Q	OK true positive fraction, TPF a            1 - β (power)	Fehler Typ 1 false positive fraction, FPF b            α	Vorhersagewert des positiven Tests PPV = a / (a + b) → a + b
	„negativ“ <b>T-</b> 1 - Q = Q'	Fehler Typ 2 false negative fraction, FNF c            β	OK true negative fraction, TNF d            1 - α	Vorhersagewert des negativen Tests NPV = d / (c + d) → c + d
Summe		a + c    ↓    Sensitivität r = a / (a + c)	b + d    ↓    Spezifität s = d / (b + d)	N

Vorhersagewerte:

$$PPV = \frac{rP}{rP + (1-s)(1-P)}$$

$$NPV = \frac{s(1-P)}{s(1-P) + (1-r)P}$$

		Status: krank / gesund		Summe
		D+ „Fälle“	D- „Kontrollen“	
		$N_{\text{cases}}$	$N_{\text{controls}}$	
Test Ausgang	„positiv“ <b>T+</b> Test level: Q	OK true positive fraction, TPF Sensitivität a $1 - \beta$ (power)	Fehler Typ 1 false positive fraction, FPF b $\alpha$	
	„negativ“ <b>T-</b> $1 - Q = Q'$	Fehler Typ 2 false negative fraction, FNF c $\beta$	OK true negative fraction, TNF Spezifität d $1 - \alpha$	
Summe		a + c            ↓            Sensitivität $r = a / (a + c)$	b + d            ↓            Spezifität $s = d / (b + d)$	N

Odds Ratio :             $OR = \frac{a \cdot d}{b \cdot c} = \frac{r \cdot s}{(1 - r) \cdot (1 - s)}$



# A systematic review of the clinical validity and clinical utility of DNA testing for hereditary haemochromatosis type 1 in at-risk populations

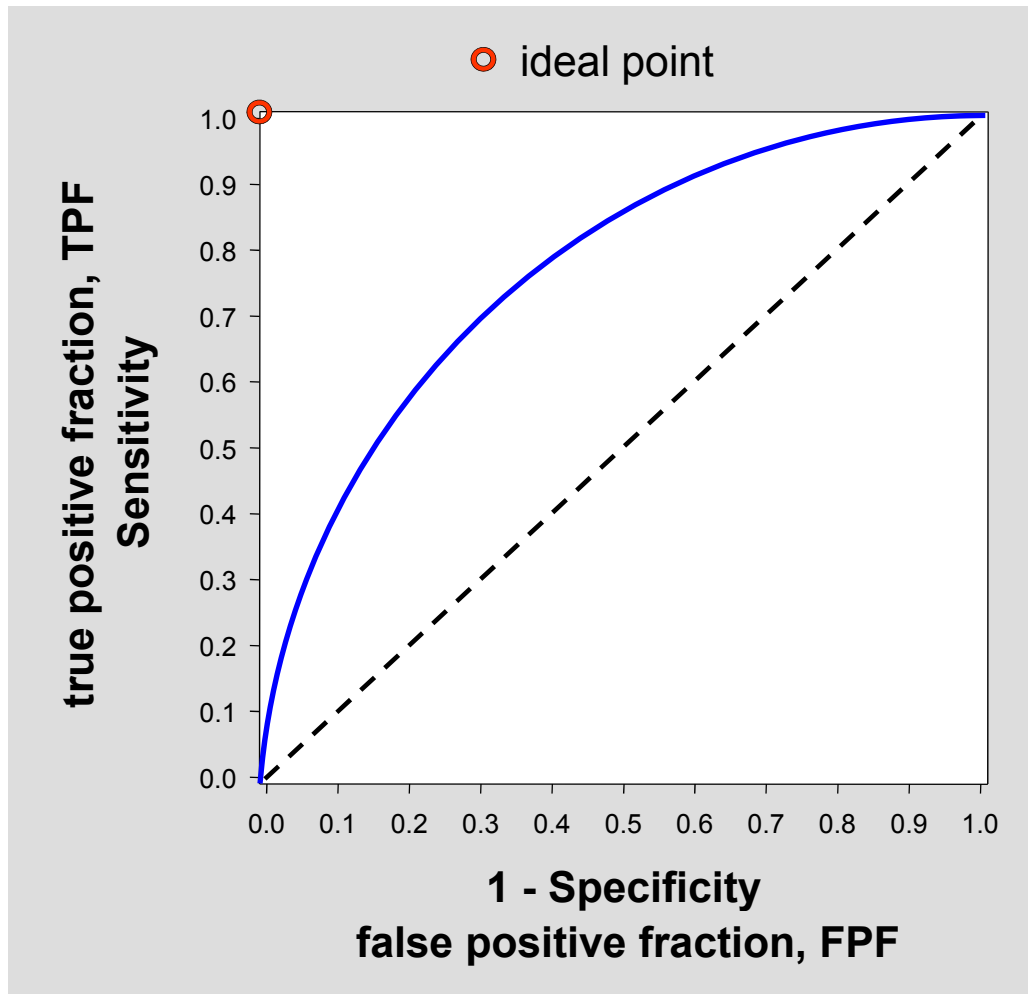
J Bryant,<sup>1</sup> K Cooper,<sup>1</sup> J Picot,<sup>1</sup> A Clegg,<sup>1</sup> P Roderick,<sup>2</sup> W Rosenberg,<sup>3</sup> C Patch<sup>2,4</sup>

J Med Genet 2008 Aug; **45**(8):513-518

**Table 2** Clinical validity of DNA testing

Study author	Sensitivity (%)	Specificity (%)	Odds Ratio DOR
Cardoso <i>et al</i> , 1998 <sup>13</sup>	92	100	∞
Hellerbrand <i>et al</i> , 2001 <sup>14</sup>	72.2	100	∞
Holmstrom <i>et al</i> , 2002 <sup>15</sup>	28.4	99.6	<b>98.7</b>
Jouanolle <i>et al</i> , 1997 <sup>16</sup>	92.4	100	∞
Mura <i>et al</i> , 2005 <sup>17</sup>	81.2	99.5	<b>860</b>
Murphy <i>et al</i> , 1998 <sup>18</sup>	90	98.8	<b>741</b>
Nielsen <i>et al</i> , 1998 <sup>19</sup>	94.6	100	∞
Ryan <i>et al</i> , 1998 <sup>20</sup>	93.3	100	∞
UK HHC Consortium, 1997 <sup>21</sup>	91.3	99	<b>1089</b>
Vantyghem <i>et al</i> , 2006 <sup>22</sup>	21.2	100	∞
Willis <i>et al</i> , 1997 <sup>23</sup>	100	99.5	∞

# Area under Curve (AUC) of the Receiver Operator Curve (ROC)



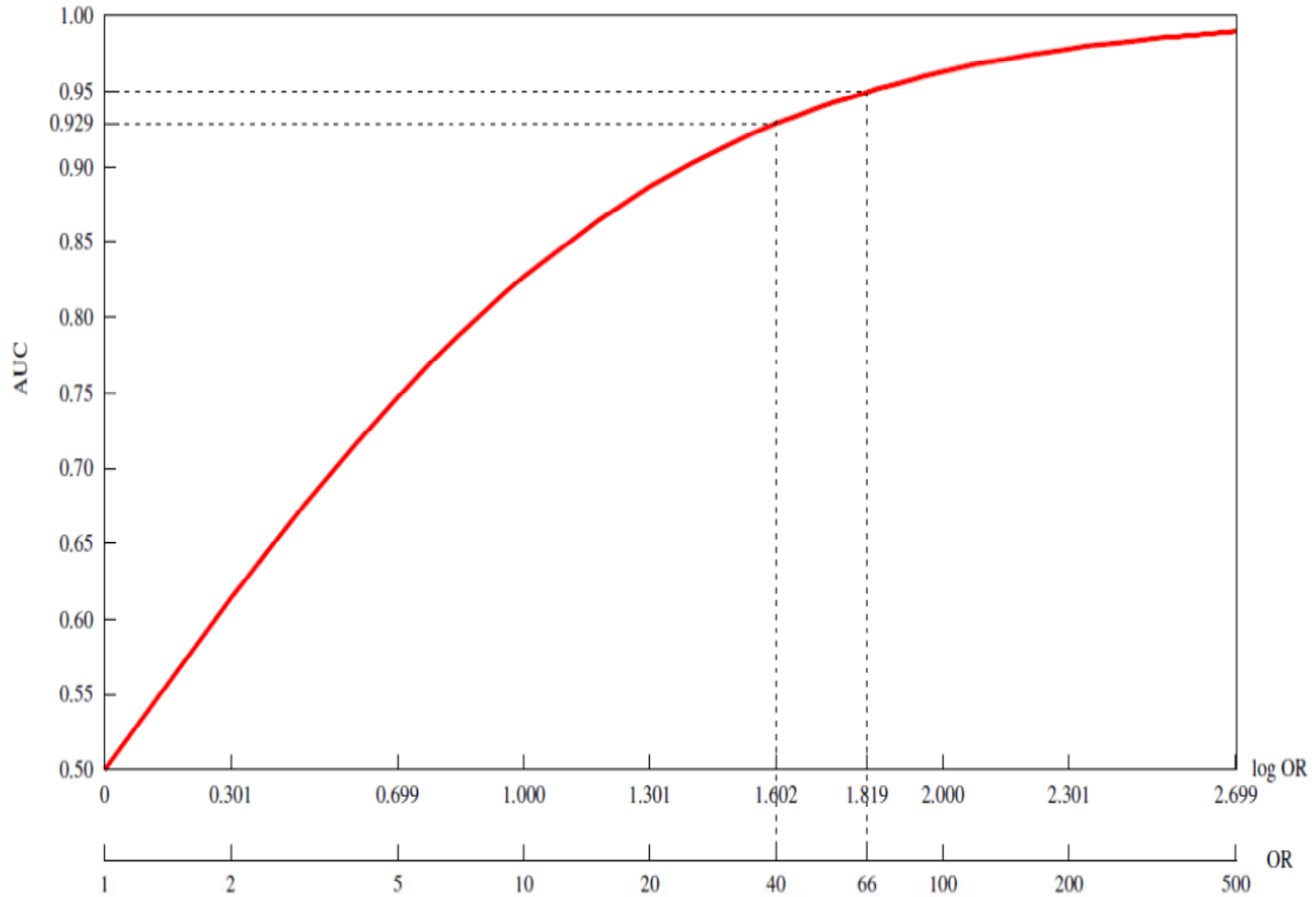
Reference AUC values:

**Tossing a coin: AUC = 0.50**

**Perfect prediction: AUC = 1.00**

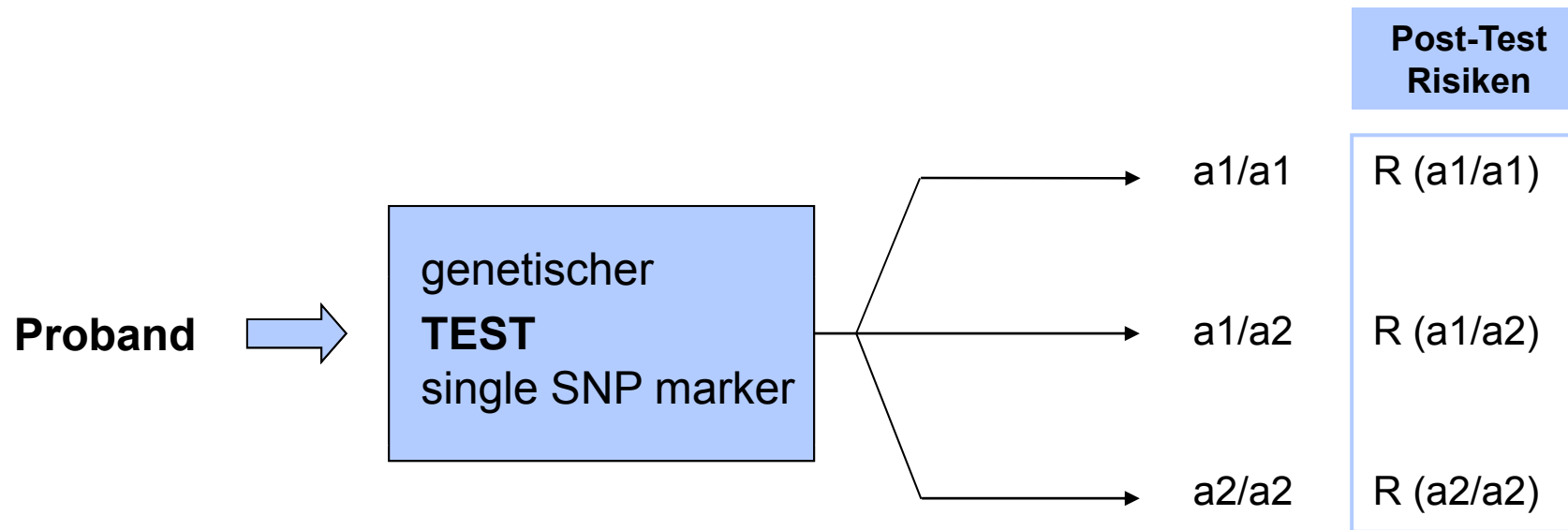
**Classical factors in CHD  
prediction: AUC ~ 0.75**

**DIAGNOSTIC ODDS RATIO ↔ AREA UNDER THE CURVE**



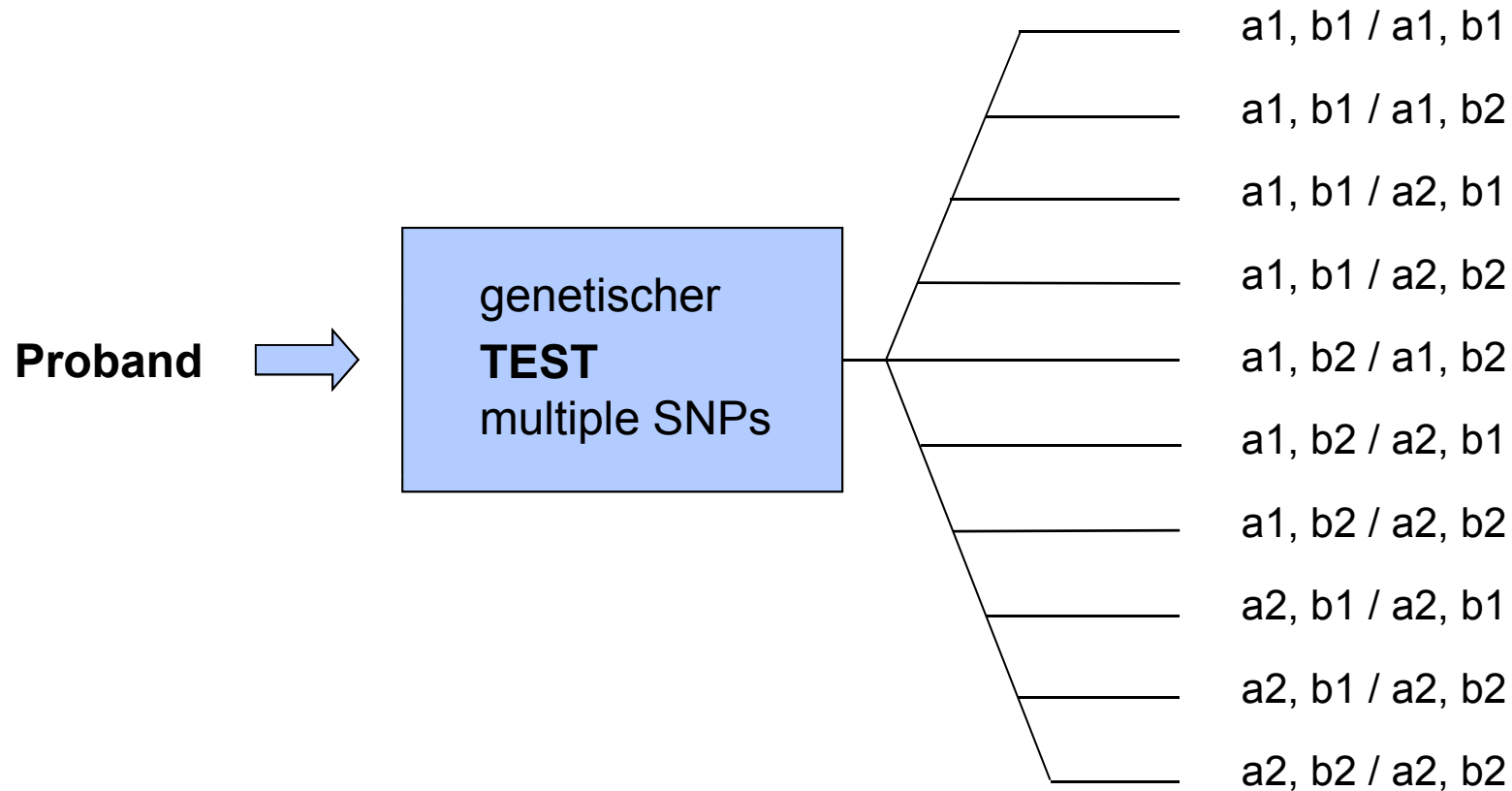
# Prädiktiver / diagnostischer Test

---



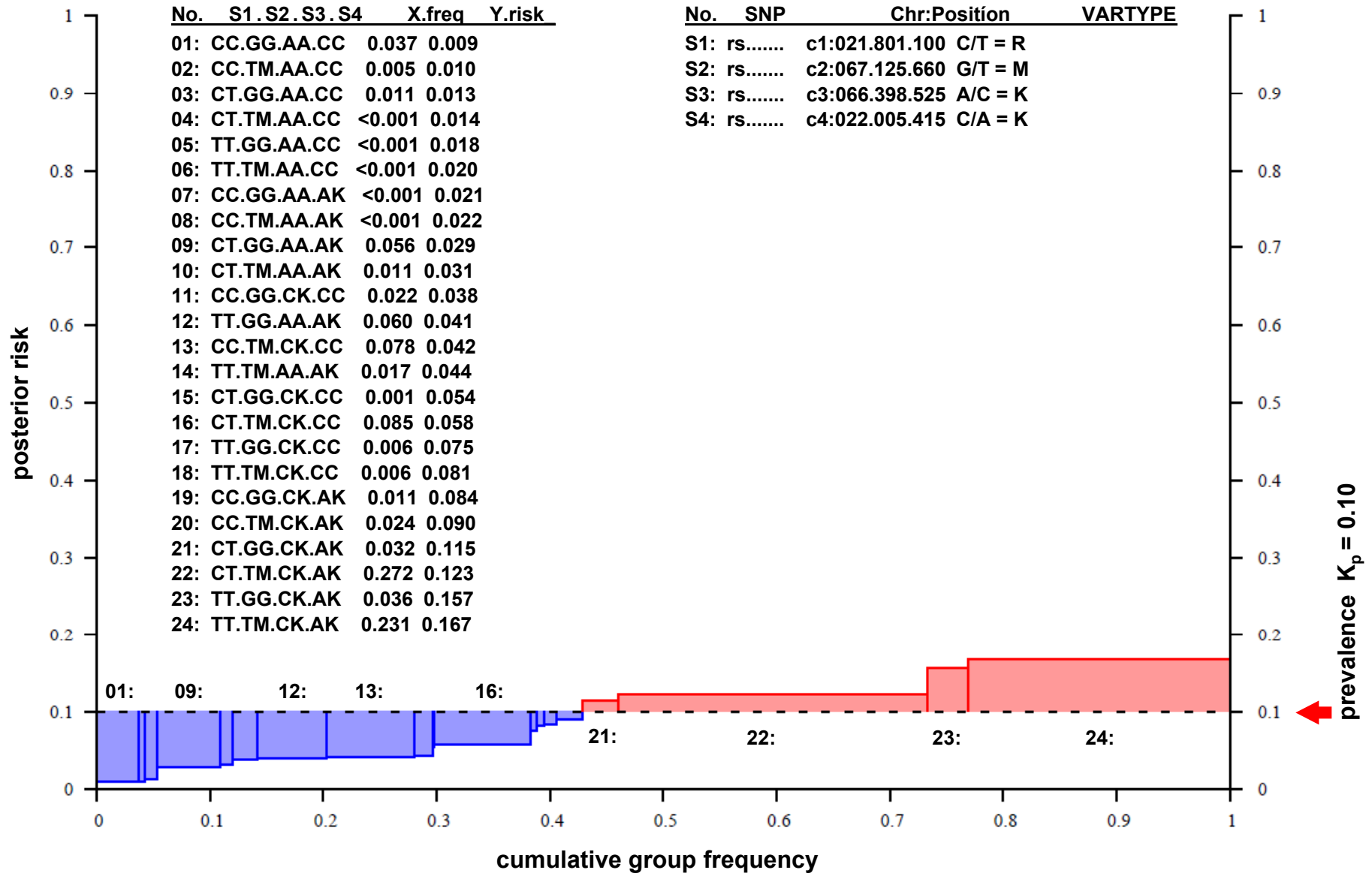
# Prädiktiver / diagnostischer Test

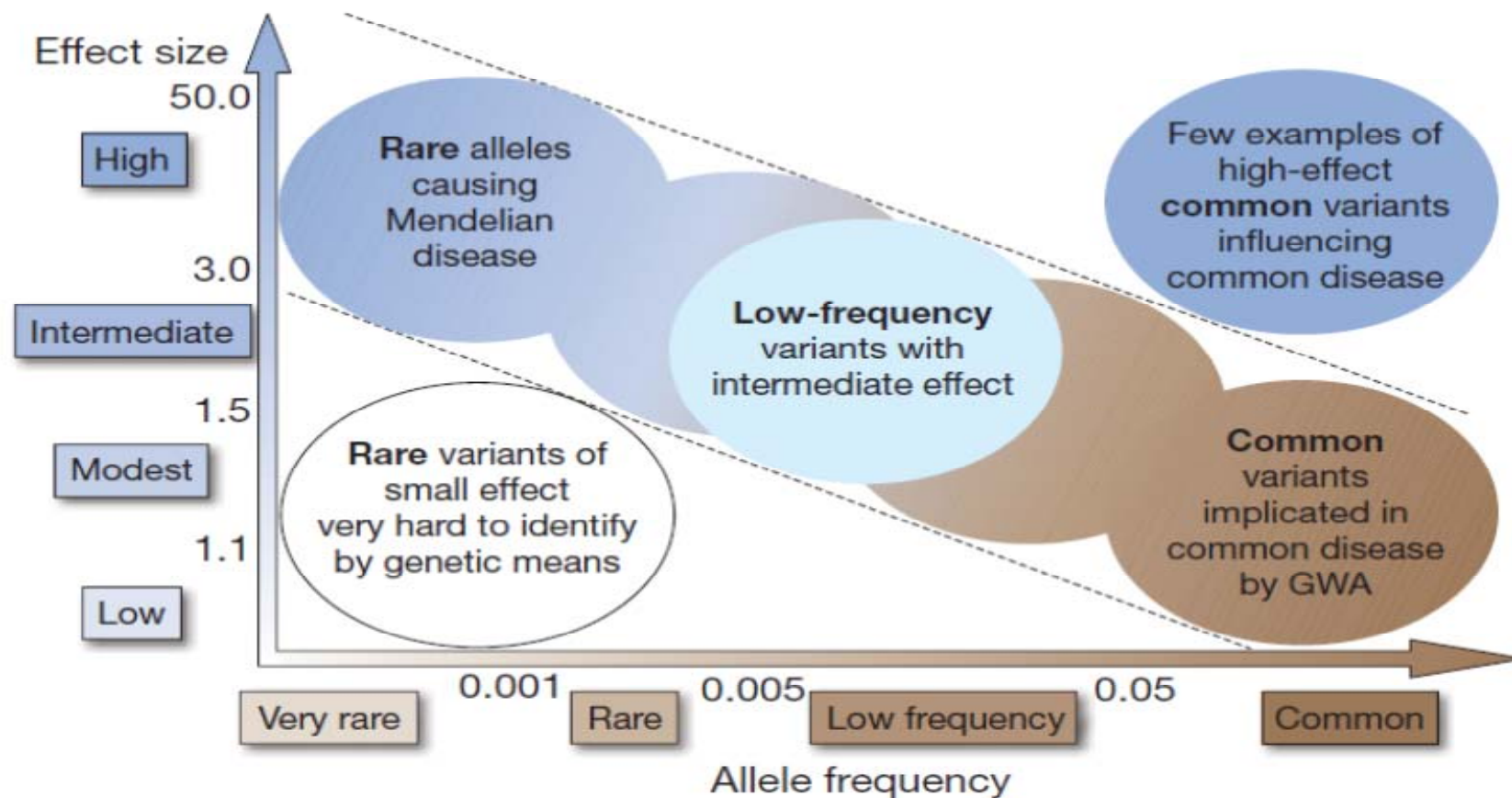
---



# Prediction Chart

risk redistribution, relative to maximum RRD = 26.8 %





**Figure 1 | Feasibility of identifying genetic variants by risk allele frequency and strength of genetic effect (odds ratio).** Most emphasis and interest lies in identifying associations with characteristics shown within diagonal dotted lines.

from: Manolio TA et al., Nature Vol. 461: 747-753 (8 October 2009)

# What if – complex traits are really complex ?

There are significant consequences when we expand the notion of causality to include the types of complex interactions that are common for biological systems:

Determinism no longer entails predictability.

Sandra D. Mitchell, *Unsimple Truths. Science, Complexity, and Policy*;  
University of Chicago Press, Chicago 2009



# The Princess on the (Mendelian) Pea



Like the famous princess, we seem to think we can always detect a mendelian pea no matter how many layers of environmental and other influences may lie over it.

Kenneth Weiss (2002):  
Goings on in Mendel's Garden,  
Evolutionary Anthropology 11: 40- 44

## Schlussfolgerungen (1)

4. Welche Bedeutung kommt der Epigenetik bei der Prädiktion komplexer Krankheiten zu?  
16. Berliner Kolloquium „Gene und Umwelt“ am 2. Mai 2012:  
*Chemische Prozesse, die sensibel auf Umwelteinflüsse reagieren, regeln wann, wo und in welchem Ausmaß Gene aktiv werden (Zitat);* Varianten der DNA-Sequenz können diese Regulation allenfalls modifizieren, und wenn, dann vermutlich nur geringfügig.
  
5. Wie schätzen Sie den Einfluss von exogenen- und Zufallsfaktoren auf die Prädiktion komplexer Krankheiten ein?  
Exogene Faktoren sind epidemiologisch schwierig zu fassen und daher unzureichend charakterisiert. Demzufolge sind auch G×E Interaktionen bisher nicht oder unzureichend charakterisiert; dies wird sich auch nicht rasch grundlegend ändern. Damit fehlen die Voraussetzungen für eine prädiktive Diagnostik

## Schlussfolgerungen (2)

6. Wie schätzen Sie die bisherige Entwicklung der Auffindung von Assoziationen zwischen genetischen Informationen und krankheitsrelevanten Symptomen ein?

Genomweite Assoziationsstudien (GWAS) sind auf absehbare Zeit für viele komplexe Krankheiten das einzige Verfahren zur Identifizierung und Charakterisierung von ursächlichen Genvarianten; es handelt sich um eine hypothesenfreies Suchverfahren. Im Lichte neuerer Befunde zur genetischen Architektur komplexer Krankheiten müssen Kandidatengen-Verfahren als nachrangig angesehen werden. Mit Ausnahme der „near-Mendelian diseases“ und pharmakogenetischen Merkmalen sind die identifizierbaren Genvarianten aber diagnostisch irrelevant.

## Schlussfolgerungen (3)

7. Was halten Sie von großen Bevölkerungsscreenings hinsichtlich GWAS (genome-wide association studies)?  
Populationsgenetische Aspekte?

Nur für sehr häufige Krankheiten mag dieser Ansatz zielführend sein, und selbst da ist der Aufwand erheblich. Für relativ seltenere Krankheiten ist der Aufwand womöglich unverträglich; dies ist sicher der Fall für die seltenen erblich bedingten Krankheiten, die überwiegend monogen bedingt sind. Hier allerdings kann eine prädiktive Diagnostik sehr wirksam greifen. Das o.g. Bevölkerungsscreening wird die dafür benötigten Kontrollen beisteuern können. Im wesentlichen sind diese Untersuchungen von überwiegend wissenschaftlichem Wert.

## Schlussfolgerungen (4)

8. Wo zeichnet sich ab, dass prädiktive Tests in Ihrem Fachbereich in Zukunft eingesetzt werden?
9. Sehen Sie den Haupteinsatzbereich genetischer Tests vor allem im prädiktiven oder differentialdiagnostischen Bereich?  
Bei den häufigen, multifaktoriell bedingten (komplexen) Krankheiten sehe ich keine Möglichkeit für wirksame prädiktive oder differentialdiagnostische Tests (von vielleicht zwei bis vier Ausnahmen abgesehen).  
Hochrelevant und wirksam sind genetische Tests aber im Bereich monogener Krankheiten; diese sind zwar jeweils selten bis sehr selten, zusammengenommen aber durchaus nicht selten. Darauf abgestellte Verfahren können die in diesem Bereich bestehenden Versorgungsdefizite mindern und im individuellen Falle zu einer wesentlichen Abkürzung der in jeder Hinsicht und für alle Beteiligten belastenden „diagnostischen Odyssee“ führen.

## Schlussfolgerungen (5)

10. Welche Voraussetzungen müssen für eine klinische Anwendung der Risikovorhersage gegeben sein?  
Der jeweilige Test muss durch ein gezieltes und mehrstufiges Verfahren analytisch und klinisch-genetisch validiert sein; entsprechende Studiendesigns wurden in den letzten Jahren analog zu den klinischen Studien zur Medikamentenzulassung entwickelt. Der Nachweis der diagnostischen und klinischen „utility“ ist integraler Bestandteil solcher Studien. Es ist zu erwarten, dass für viele (die meisten) komplexen Krankheiten dieser Nachweis nicht erbracht werden kann; insofern würde sich eine spezielle Regulierung erübrigen.
  
11. Welche Konsequenzen hinsichtlich der medizinischen Behandlung oder des Gesundheitsverhaltens sollten aus einem positiven prädiktiven Tests folgen?  
Wegen der in den meisten Fällen geringfügigen Risikomodifikation sind eingreifende oder potentiell risikobehaftete Massnahmen nicht zu vertreten.

## **Managing incidental findings and research results in genomic research involving biobanks and archived data sets**

Susan M. Wolf, JD<sup>1</sup>, Brittney N. Crock, JD<sup>1</sup>, Brian Van Ness, PhD<sup>1</sup>, Frances Lawrenz, PhD<sup>1</sup>, Jeffrey P. Kahn, PhD, MPH<sup>2</sup>, Laura M. Beskow, PhD, MPH<sup>3</sup>, Mildred K. Cho, PhD<sup>4</sup>, Michael F. Christman, PhD<sup>5</sup>, Robert C. Green, MD, MPH<sup>6,7</sup>, Ralph Hall, JD<sup>1</sup>; Judy Illes, PhD<sup>8</sup>, Moira Keane, MA<sup>1</sup>, Bartha M. Knoppers, JD, PhD<sup>9</sup>, Barbara A. Koenig, PhD<sup>10</sup>, Isaac S. Kohane, MD, PhD<sup>7</sup>, Bonnie LeRoy, MS<sup>1</sup>, Karen J. Maschke, PhD<sup>11</sup>, William McGeeveran, JD<sup>1</sup>, Pilar Ossorio, PhD, JD<sup>12</sup>, Lisa S. Parker, PhD<sup>13</sup>, Gloria M. Petersen, PhD<sup>14</sup>, Henry S. Richardson, JD, MPP, PhD<sup>15</sup>, Joan A. Scott, MS<sup>16</sup>, Sharon F. Terry, MA<sup>17</sup>, Benjamin S. Wilfond, MD<sup>18</sup> and Wendy A. Wolf, PhD<sup>19</sup>

### **INTRODUCTION**

An ongoing debate focuses on the question of whether researchers bear duties to analyze and offer back to research participants incidental findings (IFs) and individual research results (IRRs) generated in genetic and genomic research.