

Decisions on pharmacogenomic tests in the USA and Germany

Odette Wegwarth PhD,¹ Robert W. Day MD PhD MPH² and Gerd Gigerenzer PhD,³

¹Research Scientist, Max Planck Institute for Human Development, Berlin, Germany and Humboldt University, Berlin, Germany

²Director Emeritus, Fred Hutchinson Cancer Research Center, Seattle, WA, USA

³Director, Max Planck Institute for Human Development, Berlin, Germany

Keywords

diagnostics, heuristics in medicine, judgment analysis, medical decision making, oncology, pharmacogenomics

Correspondence

Odette Wegwarth
Max Planck Institute for Human
Development
Lentzeallee 94
14195 Berlin
Germany
E-mail: wegwarth@mpib-berlin.mpg.de

Accepted for publication: 27 November
2009

doi:10.1111/j.1365-2753.2010.01426.x

Abstract

Background With the increase in molecular genetic understanding of disease, diagnostic test development and availability are growing rapidly. This study investigated oncologists' decision making on using pharmacogenomic tests for cancer treatment and examined cross-cultural differences between the USA and Germany.

Methods Pilot studies revealed that the following cues play a role in decisions on pharmacogenomic tests: *stage of cancer, availability of treatment options, cost of the treatment options, severity of side effects of the treatments, therapeutic consequence of the test, cost of the test and guideline recommendation specifying use of the test*. All cues were used for designing the main study comprising nine scenarios, for each of which oncologists were asked to decide whether they would order a pharmacogenomic test.

Results On average, US oncologists opted for the test in 6.5 out of the nine scenarios (SE = 0.2), and German oncologists in 5.4 scenarios (SE = 0.2). The majority of oncologists' decisions in both the USA (76.1%) and Germany (64%) were best explained by a simple sequential model (heuristic strategy). In the USA, the information about *cost of the test* was most influential on the decisions; in Germany it was the *guideline recommendation of the test*. When the *side effects of therapy B* were described as being more severe within the scenarios, choices in favour of a non-recommended test increased by about 20% within both samples.

Conclusion Both US and German oncologists were highly inclined to use pharmacogenomic tests, but differed in what information influenced their decisions – a difference possibly explained by the differences in the health insurance systems. Although many oncologists' heuristic decisions were based on the valid cue of a test's guideline recommendation, an alarming number abandoned it when a therapy had potentially severe side effects.

Introduction

Evaluation of gene expression in cancer tumour samples now permits more accurate categorization of recurrence risk than standard clinical prognostic factors for patients with early-stage breast cancer [1–3]. One such test for breast cancer recurrence, *Oncotype DX* (Genomic Health, Redwood City, CA, USA), analyses 21 genes derived from excised tumours to predict recurrence risk for some patients with early-stage breast cancer. This test and others in development have important implications for clinicians' treatment decision making, such as whether to utilize more aggressive treatment options or not. On the other hand, for patients with low recurrence risk, clinicians may consider avoiding unnecessary treatment options and accordingly reduce potential overtreatment.

Pharmacogenomic tests have the potential to individualize prescription of therapies, which may offer the advantage of increasing patients' survival. Yet, errors such as false positive and

negative test results are drawbacks to such new technology. False positive results may result in choice of a more aggressive treatment regime for a patient who will not benefit from it but possibly suffer severe side effects, whereas a false-negative result may result in life-threatening undertreatment for a patient who would benefit from further treatment. The literature on pharmacogenomic testing is more often than not characterized by a lack of critical scrutiny of such drawbacks, and very few commentators [4] openly discuss that test inaccuracies could have a dramatic impact on patient recovery, quality of life or survival. In view of the expansion of pharmacogenomics testing, with some genomic tests already in use, a better understanding of clinicians' decision making for these tests may prove useful for not only their own understanding of issues surrounding this new technology, but also policymakers and medical educators who wish to address problems that could prevent appropriate use.

Judgment analysis has been widely used to study judgment and decision making in a variety of applied domains such as medicine [5,6]. Typically, participants are asked to make decisions on a set of cases (scenarios) that may be either real or hypothetical and that comprise a combination of cues (information). Until recently, participants' judgment policies were generally inferred by applying regression models to their decisions [7]. Although regression models have been a productive tool for investigating human decision making in applied domains, their use in describing human decision making has been questioned because these models are static and do not allow to elicit non-compensatory decision making, that is, when humans use less than complete information for their decisions [8–10]. In addition, criticism has focused on the fact that researchers using the regression approach have largely used only one – usually complex – model to describe data instead of using various models to determine which model best describes participants' decision making.

The present paper is, to the best of our knowledge, the first to describe an explorative study that investigated oncologists' decision making on pharmacogenomic tests by using accepted models of judgment analysis with different degrees of complexity. We further focused on what information is crucial for deciding whether to order these tests or not. To determine which information is important and in what manner information is processed, the following three models were applied: A *weighted additive model* – assuming that all pieces of information are considered, weighted according to their influence on the decision and then traded off against each other; an *equal-weighted model* – assuming that all pieces of information are considered and traded off against each other, but the relative influence of each on the decision is ignored; and a *simple sequential model* – assuming that only a critical amount of information, ranked and searched through by importance, is considered, and no trade-offs take place [10,11].

Decisions on pharmacogenomic testing might, however, be highly influenced by the health insurance system in which they are made. If clinicians practising under different health care systems have different expectations of what health services will be covered, their economical considerations might be as influential as clinical considerations. For instance, most people in the USA gain access to health insurance through their employers, who are not obliged to offer this benefit. Furthermore, coverage does not always expand to all health services, and losing a job may mean losing coverage. In contrast, Germany's health insurance system is mainly based on a publicly financed fund, which ensures that the great majority of the population is insured, regardless of employment status, and that most health services, drugs and dental care are covered. Given that such differences in health insurance systems may indeed impact decisions on pharmacogenomic tests, we expanded the study to oncologists in both the USA and Germany.

Method

Participants

Volunteer samples of 109 US oncologists and 111 German oncologists completed the online questionnaire of the main study. No payment was made to participants. The initial invitation to participate in the study together with a link to the online study was distributed by email to oncologists of various subspecialties across

both countries. In the USA, the ASCO Directory was used to directly email oncologists. Here, 20 persons per each letter of the alphabet were contacted, and the process was repeated until the arbitrarily predetermined count of 100 respondents was met. In Germany, medical organizations¹ in the field of oncology were approached and asked to provide their members per email the link to the study and invite them to participate. Such organizations were approached until the count of 100 respondents was again met. Both samples involved oncologists of all seniorities and from all possible work sites (private practice to training hospital). The two samples were alike in the distribution of age, specialties, years since graduation and seniority, but different in the distribution of work sites. A breakdown of the respondents' characteristics is shown in Table 1.

Procedure

To develop the scenario-based questionnaire for the main study, a pilot study involving seven US and 12 German oncologists (specialized in medical/gynecological oncology) determined what information was relevant for deciding whether to utilize a pharmacogenomic test. The following cues were crucial: *stage of cancer, availability of treatment options, costs per treatment option, severity of the side effects of the treatments involved, therapeutic consequence of the test,*² *cost of the test and guideline recommendation specifying use of the test.* The first three cues were held constant and introduced all the scenarios:

Imagine that you need to decide on an adjuvant therapy for a patient with a common solid (non-hematological) tumour.³ Based on the stage, grade, age and related clinical information assume that the current standard therapy is A plus B. Based on clinical trials, however, it is known that out of 1000 of such patients, 700 so-called low-recurrence-risk patients would remain disease-free for the next 5 years when having therapy A alone. Of the remaining 300 patients who have a high risk of recurrence, an additional 50 patients remain disease-free for the next 5 years when receiving therapy B in addition to therapy A. The cost of therapy A is \$2000 per patient; the cost of B is \$10 000 per patient. Given that standard prognostic factors do not satisfyingly predict which patient was at low or high risk of recurrence, all patients received therapy A and B. However, there is now a new test that predicts with reasonable accuracy those patients with a

¹ German Network of Evidence-Based Medicine (EBM), the German Cancer Society (DKG) with their related Consortium of Internal Oncologists (AIO), the Consortium of Gynecological Oncologists (AGO), as well as the Consortium of Urological Oncologists (AUO), Working Committee of German Tumor Centers e.V. (ADT), Northeast German Society of Gynecological Oncology e.V. (NOGGO) and the German Society of Hematology and Oncology (DGHO).

² For pilot study participants, a therapeutic consequence referred to how many patients are spared unnecessary chemotherapy (true negative rate of a test), and how many people are at risk of being undertreated (false-negative rate) by taking the test. Participants were provided with this information in an easy-to-read format (see Appendix 1). Note that this format diverges from what physicians would usually have available.

³ In order to be able to involve oncologists of various specialties, the introduction was kept general, although it resembled an adjuvant breast cancer treatment.

	Germany (<i>n</i> = 111)	USA (<i>n</i> = 109)
Age (years)		
<35	<i>n</i> = 7	<i>n</i> = 11
35–50	<i>n</i> = 71	<i>n</i> = 61
>50	<i>n</i> = 33	<i>n</i> = 37
Specialization		
Open field	Medical/Internal: <i>n</i> = 83 Gynecological/urological: <i>n</i> = 28	Medical/Internal: <i>n</i> = 86 Gynecological/urological: <i>n</i> = 23
Years since graduation		
Open field	4 years to 33 years, <i>M</i> = 18 years 1 missed value	2 years to 40 years, <i>M</i> = 15.8 years
Work site*		
Training hospital	<i>n</i> = 41	<i>n</i> = 58
Hospital	<i>n</i> = 50	<i>n</i> = 19
Private practice	<i>n</i> = 19 1 missed value	<i>n</i> = 31 1 missed value
Seniority [†]		
Head physician	<i>n</i> = 25	<i>n</i> = 38
Senior physician	<i>n</i> = 49	<i>n</i> = 33
Junior physician	<i>n</i> = 17 1 missed value	<i>n</i> = 18

**P* < 0.05.

[†]Numbers for 'seniority' do not add up to the full number of participants within each sample because oncologists practising in private practices only indicated their seniority if it applied.

low risk of recurrence (700/1000) who could be spared over-treatment with therapy B and that also indicates which patients have a high risk of recurrence (50/1000) and need therapy B as well.

Following the introduction, nine scenarios were presented, for each of which oncologists were asked whether they would use the test for making a treatment decision or not. Scenarios were developed by creating different levels for *severity of the side effects of therapy B*, *therapeutic consequence of the test*, *cost of the test* and *guideline recommendation specifying use of the test*. For example, in the case of the cue *therapeutic consequences of the test*, oncologists received information on how many of the patients with low risk of recurrence (700 out of 1000) were correctly identified as such by the test (spared overtreatment with B). This information was held constant at 350 out of 700 among all scenarios. For that cue, oncologists additionally received information on how many of those with high risk of recurrence (50 out 1000) were not correctly identified as such and thus exposed to the risk of undertreatment. This information was varied among the scenarios by two, four and six out of 50 patients. Figure 1 provides an example of a scenario. Table 2 provides the levels for all four cues varied across the scenarios.

Following the nine scenarios, oncologists had to indicate whether they considered the situation and scenarios described in the questionnaire straightforward and whether important information was missing in the scenarios. At the end, they were requested to provide demographic information (shown in Table 1).

Data analysis

The three models of decision making were programmed in MATLAB (The MathWorks, Inc., 3 Apple Hill Drive, Natick, MA

Table 1 Breakdown of the respondents' characteristics of the US and German sample

01760-2098). Each oncologist's decisions were analysed by each of the three models using MATLAB in order to determine which of the models describes the decision data best. The resulting fit of each model for each oncologist's decisions was recorded and then – together with all other data – stored and further analysed with spss 16 (SPSS Inc. Headquarters, 233 S. Wacker Drive, Chicago, Illinois 60606).

Results

Decision-making strategies on pharmacogenomic tests

In general, both US oncologists and German oncologists were very willing to order pharmacogenomic tests. Whereas US oncologists, on average, opted for the test in 6.5 (SE = 0.2) out of the nine scenarios, German oncologists opted for it in 5.4 (SE = 0.2) out of the nine scenarios [mean difference: –1.08, 95% CI (–1.61, –0.55)] (Fig. 2).

The majority of oncologists' decisions in both the USA (76.1%) and Germany (64%) were best explained by the simple sequential model, which assumes that the decision maker does not take all reasons into account. For the remaining participants, either two or all three of the models provided an equally good description of their decisions (USA: 19%, Germany: 22%) or the weighted additive model (USA: 0%, Germany: 8%) and the equal-weighted model (USA: 5%, Germany: 6%) provided the best description.

For the large group of oncologists whose decisions were best described by the simple sequential model, US oncologists considered more cues on average (*M* = 1.9; SE = 0.1) for their decisions than did German oncologists (*M* = 1.5; SE = 0.1) [mean difference: –0.43, 95% CI (–0.64, –0.21)]. Whereas among the US

Information about therapy B	
Therapy B causes serious side effects DURING TREATMENT .	
Information about the test	
350 patients would be spared over-treatment with B (out of 700/1000), 2 patients would be under-treated (out of 50/1000).	
Doing the test costs €/\$800 . Covered by health insurance.	
Doing the test is RECOMMENDED by your respective medical societies and guidelines	
I would use the test: (or recommend it to the patient, respectively)	
YES <input type="checkbox"/>	NO <input type="checkbox"/>

Figure 1 Example of one of nine scenarios in which oncologists had to decide on whether or not to use a pharmacogenomic test under the described conditions.

Table 2 Cues used for the development of scenarios and their respective levels

Reasons	Level
Severity of the side effects of the treatment	(0) Serious side effects during treatment (1) Permanent serious side effects
Therapeutic consequence of the test	(0) 350 patients would be spared overtreatment (out of 700/1000); two patients would be undertreated (out of 50/1000). (1) 350 patients would be spared overtreatment (out of 700/1000); four patients would be undertreated (out of 50/1000). (2) 350 patients would be spared overtreatment (out of 700/1000); six patients would be undertreated (out of 50/1000).
Cost of the test	(0) €/\$300 (1) €/\$800 (2) €/\$1300
Guidelines recommendation of the test	(0) Neither recommended by your experts nor by your respective medical society or guidelines (1) Recommended by some of your experts, but not by your respective medical society or guidelines (2) Recommended by your respective medical society and guidelines

oncologists, 44% considered one cue, 23.9% two cues and 32.1% three cues before making their final decision on a scenario, among the German oncologists, 70.3% considered one cue, 13.5% two cues and 16.2% three cues.

Importance of information

For US oncologists, the most influential information on their decision was (1) cost of the test, followed by (2) whether the test was recommended by guidelines, (3) the risk of undertreatment associated with the test and (4) the severity of the side effects of therapy B. For German oncologists, in contrast, the most influential information was (1) whether the test was recommended by guidelines, followed by (2) the risk of undertreatment accompanying the test, (3) the cost of the test and (4) the severity of the side effects of therapy B.

Figure 3 provides an example of a decision model, which combines the most commonly applied form of one-reason decision making among German oncologists and the most influential information within this sample.

Impact of medical guidelines on decisions

Medical guideline recommendations clearly shaped oncologists' decisions in our study. For the scenarios where the test was described as being recommended by guidelines, oncologists in both countries almost always opted for ordering the test, whereas for scenarios where the test was described as not being recommended, substantially fewer oncologists opted for it (USA: M = 97.3%, SE = 0.1 vs. M = 40.3%, SE = 1.1; Germany: M = 96.3%, SE = 0.1 vs. M = 36.8%, SE = 1.3). The information on guideline recommendation influenced US and German

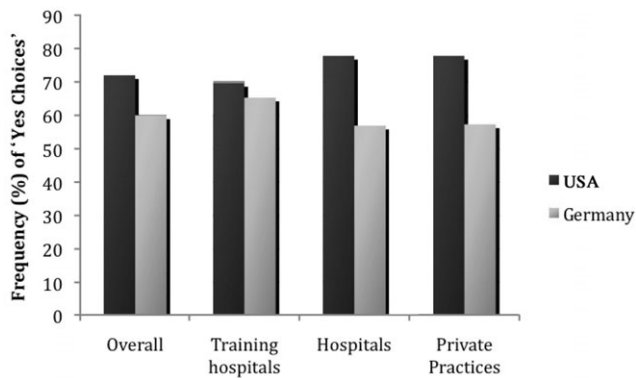


Figure 2 Frequency of 'yes choices' per country for all oncologists as well as separately for oncologists coming from training hospitals, hospitals and private practices.

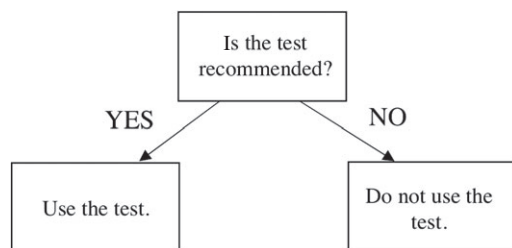


Figure 3 Example of a German one-reason decision displaying the most influential information within this sample.

oncologists in the same fashion [test recommended: $d = 0.82$, 95% CI (0.55, 1.10); test not recommended: $d = 0.14$, 95% CI (-0.12, 0.41)].

However, an increase in the severity of side effects of therapy B also led to an increase in both samples' willingness to order a non-recommended test. Considering these scenarios where the test was described as being not recommended by guidelines, on average 30.3% (SE = 0.6) of US oncologists' choices and 27% (SE = 1.0) of German oncologists' choice were in favour of such a non-recommended test when side effects of therapy B were described as being less severe (serious side effects during treatment). In contrast, when the side effects of therapy B were described as being more severe (permanent serious side effects) within scenarios, 50.3% (SE = 0.2) of the US oncologists' choices and 46.7% (SE = 0.1) of the German oncologists' choices were in favour of the non-recommended test. The information on the severity of therapy B's side effects shaped oncologists' decisions in both countries alike [side effects during treatment: $d = 0.36$, 95% CI (0.10, 0.63); permanent side effects: $d = 0.60$, 95% CI (0.33, 0.87)].

Representativeness of the questionnaire

The majority of US oncologists (74%) and German oncologists (79%) indicated that the described situation and the scenarios were straightforward. A minority of 30 US oncologists and 17 German

oncologists stated that information was missing within the scenarios that would have been important for making a decision. Asked to be more specific, the majority of these oncologists noted the absence of information on patient's preferences in having such a test (USA: 16, Germany: 5) and on the nature of the side effects of therapy B (USA: 10, Germany: 3). For the remaining German requests, one oncologist would have liked information on the Karnofsky index, and two further oncologists would have liked information on the patient's prognosis. Six out of the 17 German oncologists did not specify what other information they would have needed. For the remaining US requests, three oncologists noted the absence of information on the incidence of the disease, and one other would have liked information on the evidence of the data provided.

Conclusions

The aim of the main study was to learn about oncologists' decision-making strategies and use of information when opting for or against a pharmacogenomic test, and to further contrast effects of different health care systems.

The majority of US and German oncologists did not consider all information before arriving at their decisions on ordering a pharmacogenomic test and thus decided heuristically. This finding might be of concern to some readers. However, as demonstrated elsewhere [11–14], using less than complete information does not necessarily imply forfeiting decisional accuracy, but can lead to even more accurate decisions. In most models of rationality, it is taken for granted that the quality of decisions always improves – or at least cannot diminish – with an increasing amount of information. This assumption, however, is incorrect; the relationship between amount of information and quality of decisions is often illustrated by an inverse U-shaped curve [11,15]. Specifically, when uncertainty is high, as it is in numerous medical situations, the decision maker needs to ignore part of the available information in order to make robust predictions. For instance, in a study evaluating the ability of two different complex decision tools to correctly classify actual heart attack patients, a simple decision tree taking a maximum of three pieces of information into account was more accurate than a complex decision tool resting on some 50 pieces of information [16]. With the simple decision tree, doctors correctly assigned the largest proportion of patients who subsequently had myocardial infarction to the coronary care unit and at the same time had a comparatively low false alarm rate. In the study presented here, a recommendation of the test by medical guidelines nearly always resulted in deciding in favour of the test for both US and German oncologists, and choice behaviour differed notably for cases in which the test was not recommended. Respecting medical guidelines can diminish inappropriate practice; standardizes medical care; raises quality of care; which as a consequence will reduce several kinds of risk to the patient, the health care provider, medical insurers and health plans. The finding that many oncologists relied on less information and focused on valid cues instead – particularly in a novel situation such as that presented within our study – thus points to a reasonable behaviour.

However, we also found that guideline recommendations were sometimes abandoned by oncologists when further information was considered. On average, in 40.3% and 36.8% of the scenarios,

respectively, US and German oncologists opted for a non-recommended test. These numbers were considerably higher when more severe side effects of therapy B were involved. This observed tendency is indeed a cause for concern. Although the design of the main study⁴ did not allow for determining whether participants were able to correctly interpret probabilistic test information, the pilot study and several other studies revealed that many doctors are not able to correctly understand and calculate the statistics of test accuracy [17–20]. As a consequence, some clinicians may wrongly order tests for which benefits are not sufficiently proven.

In comparison, US and German oncologists differed mainly in the order of the information that most influenced their decisions. National differences in what information most influenced decisions on testing may reveal differences between two different health care systems. Costs of the test ranked only third for German oncologists, who practise in a health care system where almost everybody is insured and coverage of health services is clearly delineated by the government. The same information ranked first for the US oncologists, who may more often experience financial restrictions in a system in which a considerable number of patients are not insured and health service coverage depends on several aspects, including the socioeconomic status of the patient.

A limitation of the study is that participants were not a randomly assembled group of oncologists. Thus, results might be prone to selection bias resulting in an overestimation or underestimation of their readiness to order pharmacogenomic tests as well as of the order of the importance of offered information. However, all participants are involved in deciding on the course of cancer therapy on a daily basis and thus belong to the target group for whom such a test may become of relevance. Furthermore, both samples included oncologists from all possible work sites and of a variety of ages, seniorities and related characteristics such that we assume a fair generalizability of our findings.

Our study revealed a high interest on the part of both US and German oncologists in having a test at hand that could provide more rationale in a field where consensus on the best care does not always exist and therapy options can result in life-threatening consequences. However, oncologists' tendency to also opt for non-recommended tests without perhaps being fully able to appreciate the tests' accuracy is an issue that needs to be tackled in medical education and continuing medical training before a wider range of pharmacogenomic tests becomes available to doctors and their patients.

Authors contribution

O. Wegwarth has full access to all of the data in the study and takes responsibility for the integrity of the data analysis. Study concept and design were developed by O. Wegwarth and G. Gigerenzer. Data were acquired by O. Wegwarth with support by R. Day.

⁴ Because the aim of the main study was to investigate oncologists' decision making, not to test their knowledge of test theory, participants were provided with information in the scenarios on the true negative and false-negative rates in an easy-to-read format, that is, conditional probabilities for the true negative and false-negative rates were translated into simple frequency statements (see Table 2). Note that this format diverges from what physicians would find in their customary setting, where they may have difficulty in understanding test statistics and may regard the information differently than in this study.

Analysis and interpretation of the data were performed by O. Wegwarth with the assistance of G. Gigerenzer. The manuscript was drafted by O. Wegwarth and critically revised by R. Day and G. Gigerenzer.

Ethical approval

The study was approved by the Ethics Committee of the Max Planck Institute for Human Development. All participants consented to participation through an online consent form at the beginning of the survey.

Conflict of interests and funding

The reported study was conducted as a part of O. Wegwarth's Ph.D. project at the Humboldt University, Berlin (Germany). The Ph.D. was funded by a grant from Epigenomics AG, Berlin. The funding organization had no role in the design or conduct of the study, the collection, management, analysis or interpretation of data, or the preparation and approval of the submitted manuscript. Dr Robert Day, President and Director Emeritus of the Fred Hutchinson Cancer Center Seattle, is also chief medical officer at the Epigenomics branch in Seattle (WA). No further potential conflict of interests are declared.

Acknowledgements

We thank the 220 'nameless' oncologists of the main study and the 19 oncologists of the pilot study for making this study possible.

References

1. Buyse, M., Loi, S., van't Veer, L., Viale, G., Delorenzi, M., Glas, A. M., *et al.* (2006) Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. *Journal of the National Cancer Institute*, 98 (17), 1183–1192.
2. Foekens, J. A., Atkins, D., Zhang, Y., Sweep, F. C., Harbeck, N., Paradiso, A., *et al.* (2006) Multicenter validation of a gene expression-based prognostic signature in lymph node-negative primary breast cancer. *Journal of Clinical Oncology*, 24 (11), 1665–1671.
3. Paik, S., Shak, S., Tang, G., Kim, C., Baker, J., Cronin, M., *et al.* (2004) A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *New England Journal of Medicine*, 351, 2817–2826.
4. Holtzman, N. A. (2003) Clinical utility of pharmacogenetics and pharmacogenomics. In *Pharmacogenomics: social, ethical, and clinical dimensions* (ed. M. A. Rothstein), New Jersey: John Wiley & Sons, Inc.
5. Wigton, R. S., Hoellerich, V. L. & Patil, K. D. (1986) How physicians use clinical information in diagnosing pulmonary embolism: an application of conjoint analysis. *Medical Decision Making*, 6, 2–11.
6. Wigton, R. S. (1987) The use of computer simulation in teaching clinical diagnosis. *Computer Methods and Programs in Biomedicine*, 25, 111–114.
7. Cooksey, R. W. (1996) *Judgment Analysis: theory, methods, and applications*. London: Academic Press.
8. Dhami, M. K. & Harries, C. (2001) Fast and frugal versus regression models of human judgment. *Thinking & Reasoning*, 7, 5–27.
9. Gigerenzer, G. & Goldstein, D. G. (1996) Reasoning the fast and frugal way: models of bounded rationality. *Psychological Review*, 103, 650–669.

10. Dhami, M. K. & Ayton, P. (2001) Bailing and jailing the fast and frugal way. *Journal of Behavioral Decision Making*, 14, 141–168.
11. Gigerenzer, G., Todd, P. M. & the ABC Research Group (1999) *Simple Heuristics That Make Us Smart*. New York: Oxford University Press.
12. Wegwarth, O., Gaissmaier, W. & Gigerenzer, G. (2009) Smart strategies for doctors and doctors in training: heuristics in medicine. *Medical Education*, 43 (8), 721–728.
13. Martignon, L., Katsikopoulos, K. V. & Woike, J. K. (2008) Categorization with limited resources: a family of simple heuristics. *Journal of Mathematical Psychology*, 52 (6), 352–361.
14. Brighton, H. (2006) Robust inference with simple cognitive models. In *Between A Rock and A Hard Place: Cognitive Science Principles Meet AI-Hard Problems Papers from the AAAI Spring Symposium*, AAAI Tech. Rep. No. SS-06-03 (eds C. Lebiere & R. Wray), pp. 17–22. Menlo Park, CA: AAAI Press.
15. Pitt, M. A., Myung, I. J. & Zhang, S. (2002) Toward a method of selecting among computational models of cognition. *Psychological Review*, 109, 472–491.
16. Green, L. A. & Mehr, D. R. (1997) What alters physicians' decisions to admit to the coronary care unit? *The Journal of Family Practice*, 45, 219–226.
17. Eddy, D. M. (1982) Probabilistic reasoning in clinical medicine: problems and opportunities. In *Judgment under Uncertainty: heuristics and biases* (eds D. Kahneman, P. Slovic & A. Tversky), pp. 249–267. Cambridge: Cambridge University Press.
18. Hoffrage, U. & Gigerenzer, G. (1995) How to improve Bayesian reasoning without instruction: frequency formats. *Psychological Review*, 102, 684–704.
19. Casscells, W., Schoenberger, A. & Grayboys, T. (1978) Interpretation by physicians of clinical laboratory results. *New England Journal of Medicine*, 299, 999–1000.
20. Gigerenzer, G., Gaissmaier, W., Kurz-Milcke, E., Schwartz, L. M. & Woloshin, S. (2007) Helping doctors and patients to make sense of health statistics. *Psychological Science in the Public Interest*, 8, 53–96.

Appendix 1

