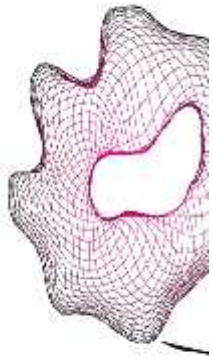
UNIVERSITY OF TWENTE.

Elicitation of stakeholder preferences in early models for Health Technology Assessment

Maarten J. IJzerman, PhD
Dept. Health Technology & Services Research (HTSR)
MIRA institute for Biomedical Technology & Technical Medicine

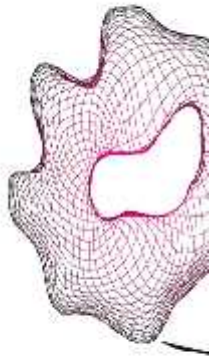
Dean Health & Biomedical Technology
School of Science & Technology

Malta, COST action – Expert Judgment in Healthcare
October, 8 2015



Background

- Healthcare: EU (-1 / 2%) vs. non-EU (3 / 4%)
  - 1 trillion US$ (USA), 400 bUS$ (China), 450 bUS$ (EU)
  - MedTech 440 bUS$ (+4%) vs. Pharma 857 bUS$ (+2,5%)
- Nearly US$ 269 billion spent on research in 2012
  - About 40% is from public sources
- A significant number of technologies does not make it to the market
- Developmental uncertainties
  - Technical (performance, specs etc.) and clinical (health need)
- Market uncertainties
  - Does it add value compared to other technologies
  - Are people willing to use the technology
  - Will there be a fair reimbursement



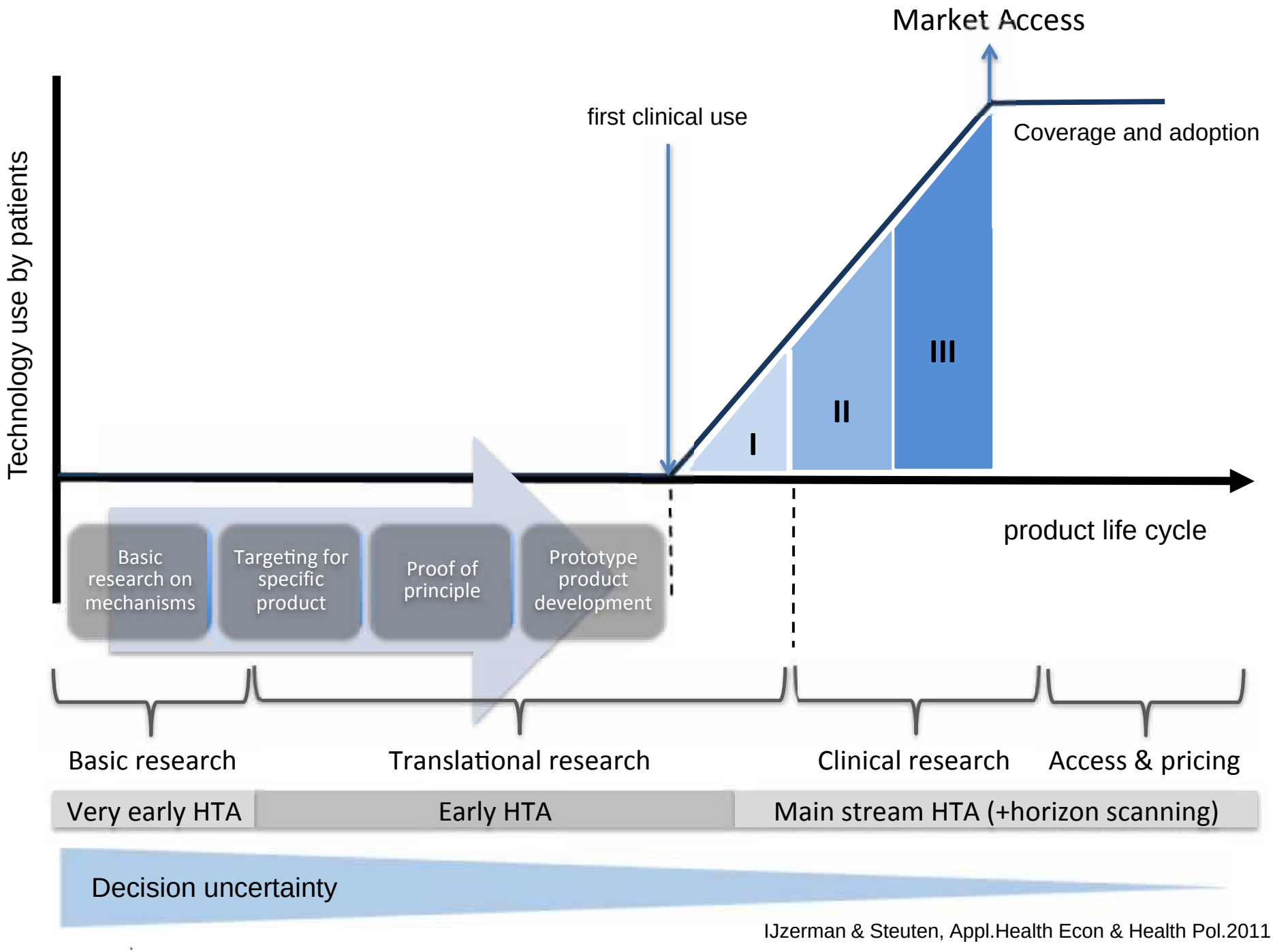
Technology increases healthcare cost

- The implementation of new medical technology accounts for between 38 and 65% of health care spending increases
  www.aetna.com/health-reform-connection/aetnas-vision/facts-about-costs.html

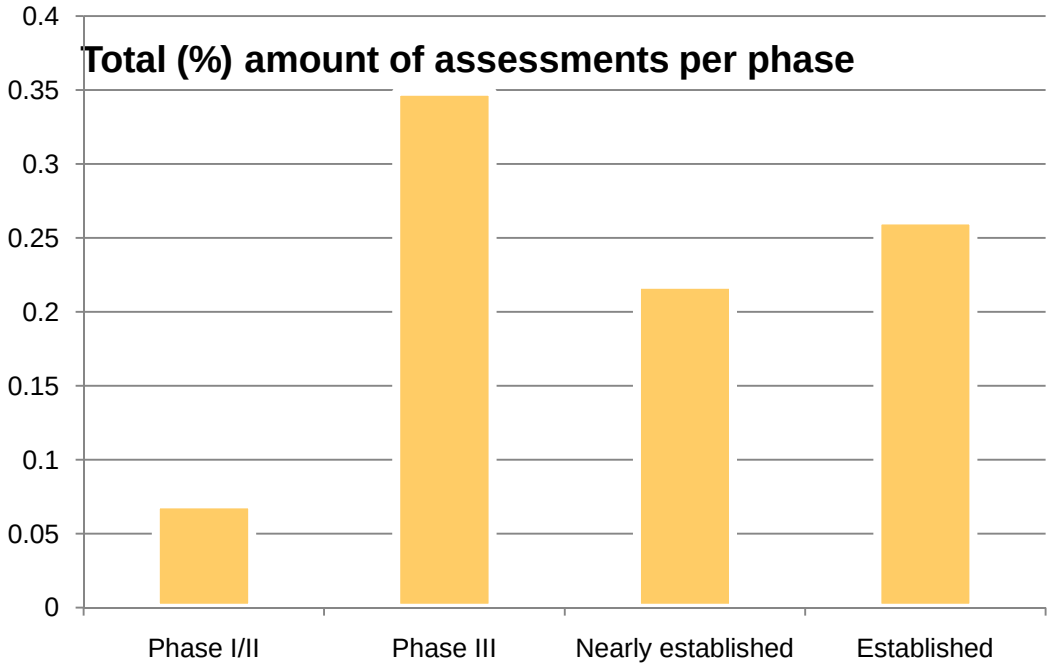
- Studies summarized in WHO's work for the Czech EU Presidency Conference on Financial Sustainability suggest that technological development accounts for between 50 and 75% of the growth in health care costs (WHO, April 2011)

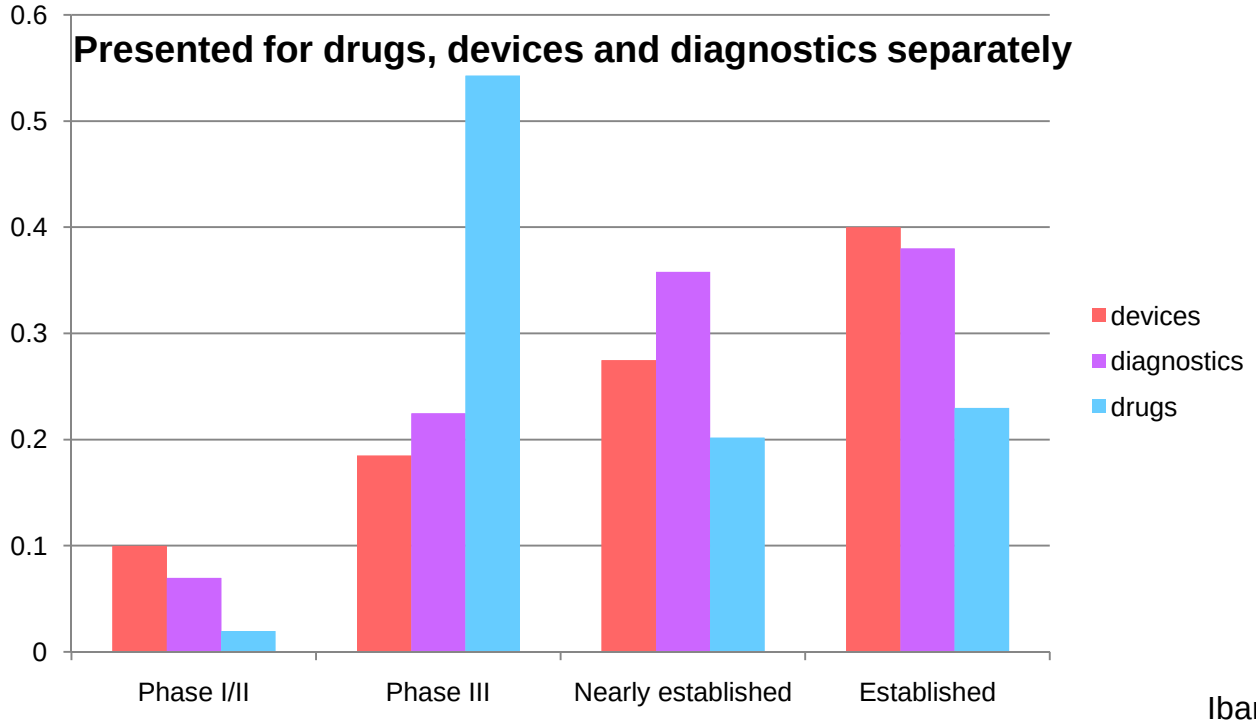
- In every industry but one, technology makes things better and cheaper. Why is it that innovation increases the cost of healthcare? MIT technology review
  www.technologyreview.com/news/518876/the-costly-paradox -of-health-care-technology/

- There is consensus among experts that technology is the most important driver of healthcare spending increases over time.
  www.forbes.com/sites/realspin/2013/04/03/whos-to-blame-for-our-rising-healthcare-costs/



IJzerman & Steuten, Appl.Health Econ & Health Pol.2011





Source:
Euroscan database

N=1085 cases

14 HTA agencies

Ibargoyen *et al*, Int. J. TA Healthcare 2009

Early health technology assessment

Table 1. Similarities and Differences between Classical HTA and Early HTA

	Classical HTA	Early HTA
Aim	Assess safety, effectiveness, and cost-effectiveness profiles of a new technology	Assess (likely) safety, effectiveness, and cost-effectiveness profiles of a new technology
Decision support	Decision support for **regulators, payers, and patients** about *market clearance, payment, and usage of a technology*	Decision-support for **manufacturers and investors** about *design and management of a technology, as well as regulatory and reimbursement strategy*
Available evidence	Usually evidence from clinical studies performed with the new technology	Evidence from early bench and animal testing, early clinical experience, and from previous generations of the technology
Influence on technology performance	Limited or no influence on clinical performance of a new technology	Potentially significant influence on (future) clinical performance of a new technology

Pietzsch & Paté-Cornell, Int. J. Techn. Assessm. Healthcare, 2008

Translational Health Economics

- How can the value of a potential new medical technology be established to inform the initiation of a translation process?
- How can evidence of an emerging health technology's value be established as early as possible during the translation process?
- How can the translational value chain be optimized, to overcome the multiple barriers along the way to market access and reimbursement

From: Rogowski, John & IJzerman: Worldbook of Health Economics, 2016

“Expert judgment”

And suddenly, you realize how often experts are involved in our research. By asking them to value health services, to prioritize medical technologies and to determine risk tolerance and decision trade-offs.

This presentation

1. Judgment: ‘probability’ or ‘preference’ elicitation
   - **Choice-modeling**: discrete-choice and best-worse scaling
     - CRC screening: patient preferences for alternative screening
     - *ISPOR Taskforces: Bridges, 2011, Johnson, 2013, Hauber, 2015*
   - **MCDA**: weighing, ranking and prioritizing alternatives
     - Photoacoustic imaging; prioritizing further development
     - *ISPOR Taskforces MCDA: www.ispor.org*
   - **Probability elicitation**: elicitation of priors to populate HE models
     - Mammography for breast cancer screening and diagnostics
     - *Johnson et al, 2010 and Butler et al, 2015*
2. Role of expert judgment in personalized healthcare
   - Health systems approach: where to add value?
   - Utility of diagnostic information

Groothuis-Oudshoorn et al. BMC Medical Informatics and Decision Making 2014, 14:18
http://www.biomedcentral.com/1472-6947/14/18

BMC
Medical Informatics & Decision Making

RESEARCH ARTICLE Open Access

Public stated preferences and predicted uptake for genome-based colorectal cancer screening

Catharina GM Groothuis-Oudshoorn[1*], Jilles M Fermont[1,2], Janine A van Til[1] and Maarten J IJzerman[1]

Abstract

Background: Emerging developments in nanomedicine allow the development of genome-based technologies for non-invasive and individualised screening for diseases such as colorectal cancer. The main objective of this study was to measure user preferences for colorectal cancer screening using a nanopill.

Methods: A discrete choice experiment was used to estimate the preferences for five competing diagnostic techniques including the nanopill and iFOBT. Alternative screening scenarios were described using five attributes namely: preparation involved, sensitivity, specificity, complication rate and testing frequency. Fourteen random and two fixed choice tasks, each consisting of three alternatives, were offered to 2225 individuals. Data were analysed using the McFadden conditional logit model.

Results: Thirteen hundred and fifty-six respondents completed the questionnaire. The most important attributes (and preferred levels) were the screening technique (nanopill), sensitivity (100%) and preparation (no preparation). Stated screening uptake for the nanopill was 79%, compared to 76% for iFOBT. In the case of screening with the nanopill, the percentage of people preferring not to be screened would be reduced from 19.2% (iFOBT) to 16.7%.

Conclusions: Although the expected benefits of nanotechnology based colorectal cancer screening are improved screening uptake, assuming more accurate test results and less preparation involved, the relative preference of the nanopill is only slightly higher than the iFOBT. Estimating user preferences during the development of diagnostic technologies could be used to identify relative performance, including perceived benefits and harms compared to competitors allowing for significant changes to be made throughout the process of development.

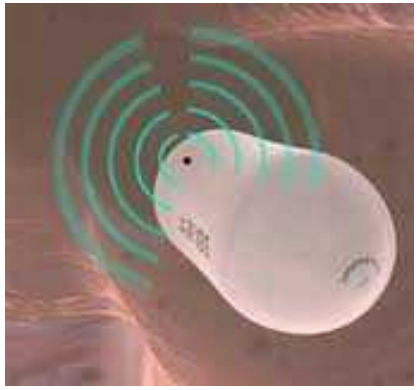
Keywords: Discrete choice experiment, Conjoint analysis, Nanopill, Colorectal cancer screening, Health technology assessment

Developments in nano-oncology for cancer detection

- Early detection of cancer in body fluids
  - Methylated DNA (Mikeska & Craig, 2014)
    - In urine (bladder cancer), gastro-intestinal tract (GIST), cervical cancer (HPV), and other
- Liquid biopsies for prognosis and to evaluate probability of distant metastasis (Kidess & Jeffrey, 2013)
  - Cell-free DNA fragments (cf-DNA)
  - Circulating Tumor Cells (CTCs)
    - CTC count (>5 CTC/7,5mL) have prognostic value for survival in BC, CRC and PC (Doyle et al, 2010)

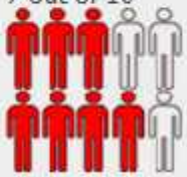
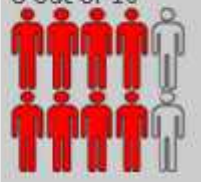
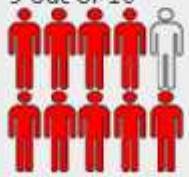
Mikeska T, Craig JM. DNA methylation biomarkers: cancer and beyond. Genes (Basel). 2014;5(3):821–64.
Kidess E, Jeffrey SS: Circulating tumor cells versus tumor-derived cell-free DNA: rivals or partners in cancer care in the era of single-cell analysis? Genome Medicine 2013, 5:70.
Miller MC, Doyle GV, Terstappen LWMM. Significance of Circulating Tumor Cells Detected by the CellSearch System in Patients with Metastatic Breast Colorectal and Prostate Cancer. J Oncol. 2010;2010:617421.

An application to new CRC screening technologies



Groothuis *et al*, BMC Med Inf & Dec Mak, 2014

Imagine that you can choose how you will be screened for colorectal cancer. Please look at the screeningtests below and select the test you prefer by clicking the button below this test.

How do you need to prepare?	Before the test you need to take laxatives which cause diarrhoea to empty your colon.	Before the test you need to take enemas which cause diarrhoea to empty your colon.	For 3 days you need to alter your diet and medication. Before the test you need to take laxatives which cause diarrhoea to empty your colon.
How is the test done?	A short flexible tube with a small camera is inserted through the anus into the last part of the colon. This test is done at a hospital.	A long flexible tube with a small camera is inserted through the anus into the full colon. During the examination you wil be sedated. This test is done at a hospital.	You need to swallow a pill that leaves your body through faeces after several hours. Your test results are wirelessly sent to your physician. This test is done at home.
How many out of 10 people <u>with</u> cancer, would the test correctly identify?	7 out of 10 	8 out of 10 	9 out of 10 
How many out of 10 people <u>without</u> cancer, would the test correctly identify?	7 out of 10 	10 out of 10 	9 out of 10 
How many out of 10.000 people who take this test have a complication?	None	10 out of 10.000	10 out of 10.000
How often do you need to take the test?	Every 5 years	Every year	Every 10 years

If you could choose between the test you chose or not to be screened for colorectal cancer, what would you prefer?

- I would still prefer the test I chose above
- I would prefer not to be screened

Technique
Sensitivity
Preparation
Frequency
Specificity
Complications
0.0
0.2
0.4
0.6
0.8
1.0
pill
stool
100%
90%
80%
70%
100%
90%
80%
70%
0%
0.01%
0.1%
1%
Pill
Stool
Short tube
tube with sedation
100%
90%
80%
70%
No preparation
Laxatives
Enemas
Diet plus laxatives
Biennial
Every 5 years
Annual
Every 10 years
100%
90%
80%
70%
None
1/10.000
10/10.000
100/10.000

Medical Devices: Evidence and Research

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Assessment of the added value of the Twente Photoacoustic Mammoscope in breast cancer diagnosis

This article was published in the following Dove Press journal:
Medical Devices: Evidence and Research
27 July 2011
Number of times this article has been viewed

Marjolein P Hilgerink[1]
Marjan JM Hummel[2]
Srirang Manohar[3]
Simon R Vaartjes[1]
Maarten J IJzerman[2]

[1]Department of Medical Physics, Medisch Spectrum Twente, Enschede, The Netherlands; [2]Health Technology and Services Research, [3]Biomedical Photonic Imaging, MIRA Institute, University of Twente, Enschede, The Netherlands

Purpose: Photoacoustic (PA) imaging is a recently developed breast cancer imaging technique. In order to enhance successful clinical implementation, we quantified the potential clinical value of different scenarios incorporating PA imaging by means of multi-criteria analysis. From this analysis, the most promising area of application for PA imaging in breast cancer diagnosis is determined, and recommendations are provided to optimize the design of PA imaging.

Methods: The added value of PA imaging was assessed in two areas of application in the diagnostic track. These areas include PA imaging as an alternative to x-ray mammography and ultrasonography in early stage diagnosis, and PA imaging as an alternative to Magnetic Resonance Imaging (MRI) in later stage diagnosis. The added value of PA imaging was assessed with respect to four main criteria (costs, diagnostic performance, patient comfort and risks). An expert panel composed of medical, technical and management experts was asked to assess the relative importance of the criteria in comparing the alternative diagnostic devices. The judgments of the experts were quantified based on the validated pairwise comparison technique of the Analytic Hierarchy Process, a technique for multi-criteria analysis. Sensitivity analysis was applied to account for the uncertainty of the outcomes.

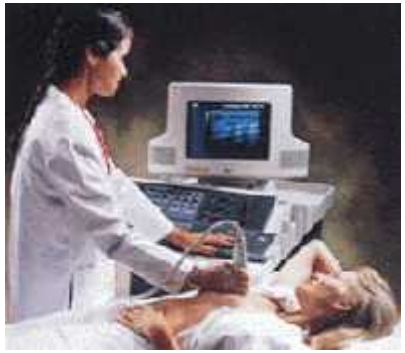
Results: Among the considered alternatives, PA imaging is the preferred technique due to its non-invasiveness, low cost and low risks. However, the experts do not expect large differences in diagnostic performance. The outcomes suggest that design changes to improve the diagnostic performance of PA imaging should focus on the quality of the reconstruction algorithm, detector sensitivity, detector bandwidth and the number of wavelengths used.

Conclusion: The AHP method was useful in recommending the most promising area of application in the diagnostic track for which PA imaging can be implemented, this being early diagnosis, as a substitute for the combined use of x-ray mammography and ultrasonography.

Keywords: technology assessment, breast cancer, diagnostic imaging, biomedical engineering

Photoacoustic imaging in breast cancer

- Mammography
  - Screening
  - **Diagnostic**
- **Ultrasound**
- **MRI**
- Biopsy



1

3

2

Hilgerink et al. Medical Devices: Evidence & Research, 2011

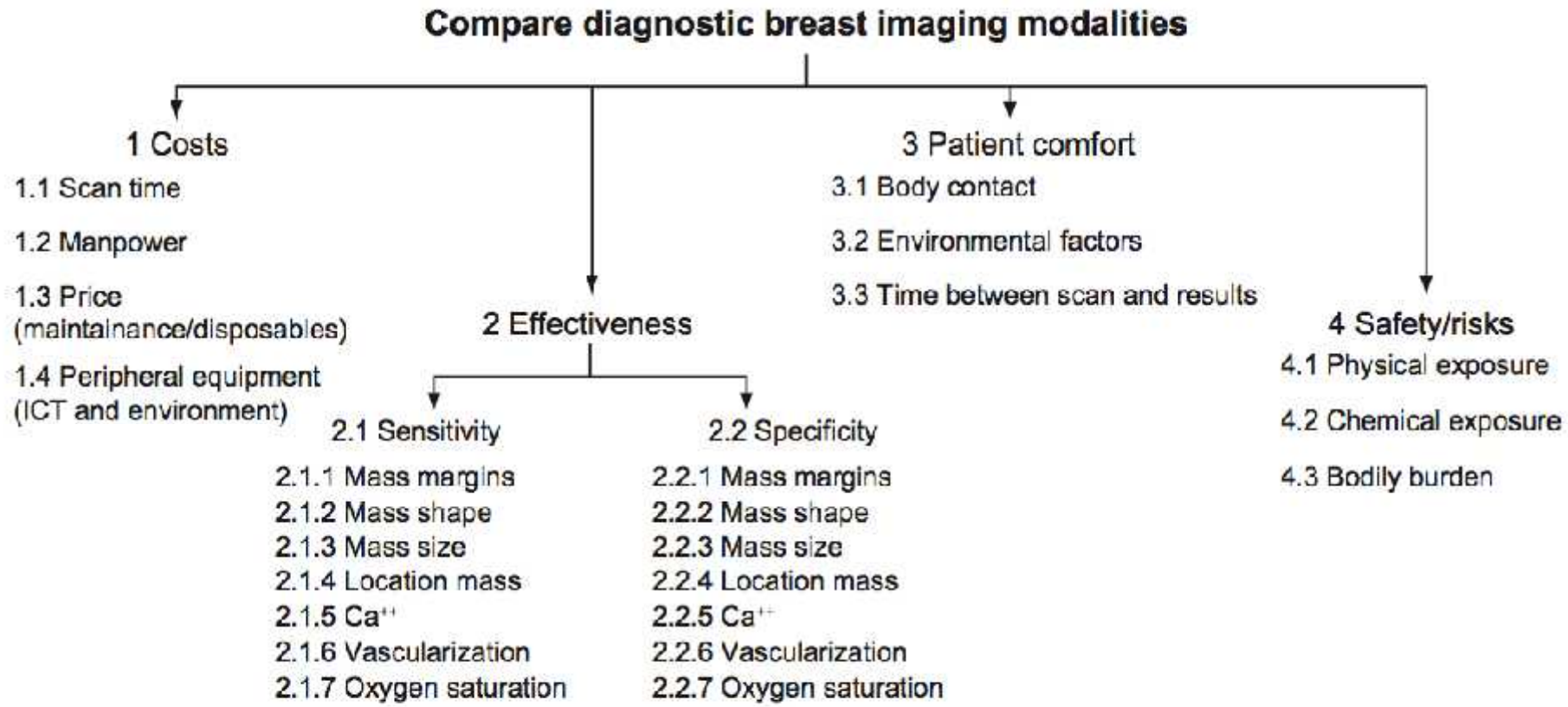
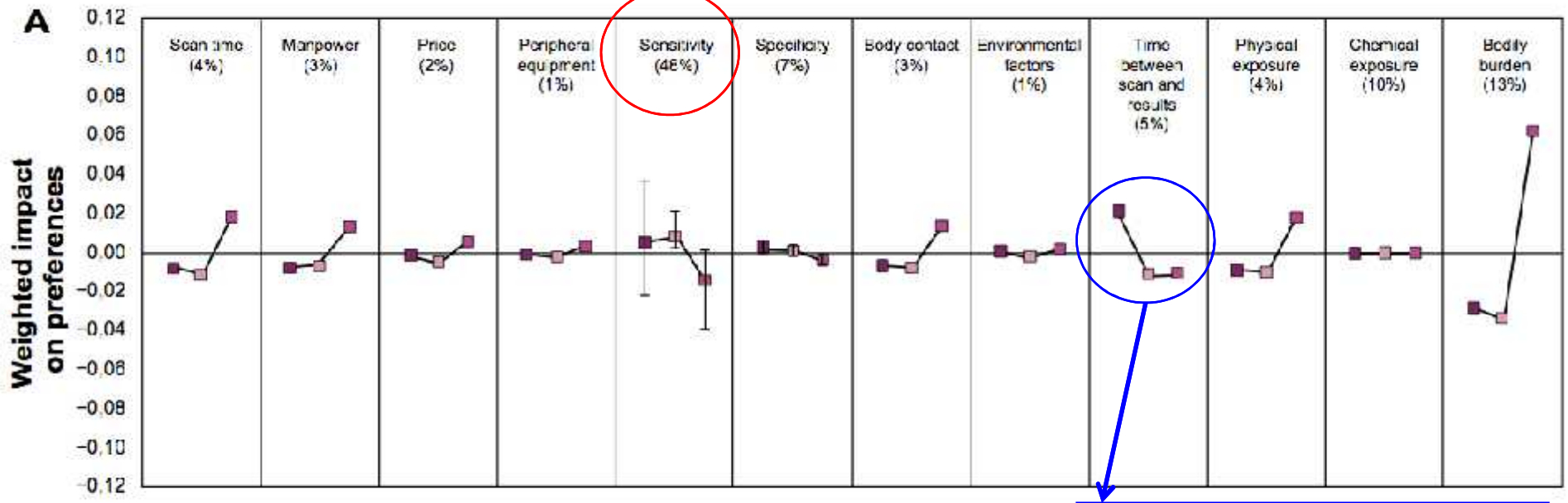


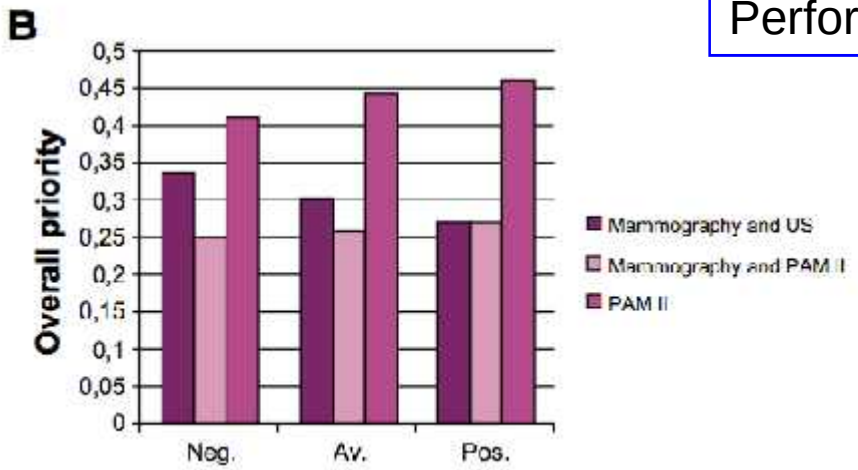
Figure 1 AHP hierarchical structure.
Abbreviation: AHP, analytic hierarchy process.

Early diagnosis, comparison with US and mammography

Contribution to overall priority



Performance on criterion



3 scenario's based on experts' uncertainty:

- Negative
- Average
- Positive

Judgments are penalized based on their uncertainty (certain, moderate, uncertain)

Appl Health Econ Health Policy
DOI 10.1007/s40258-014-0092-y

ORIGINAL RESEARCH ARTICLE

Belief Elicitation to Populate Health Economic Models of Medical Diagnostic Devices in Development

Wieke Haakma · Lotte M. G. Steuten · Laura Bojke · Maarten J. IJzerman

Abstract

Background and Objective Bayesian methods can be used to elicit experts' beliefs about the clinical value of healthcare technologies. This study investigates a belief–elicitation method for estimating diagnostic performance in an early stage of development of photoacoustic mammography (PAM) imaging versus magnetic resonance imaging (MRI) for detecting breast cancer.

Research Design Eighteen experienced radiologists ranked tumor characteristics regarding their importance to detect malignancies. With reference to MRI, radiologists estimated the true positives and negatives of PAM using the variable interval method. An overall probability density function was determined using linear opinion pooling, weighted for individual experts' experience.

Result The most important tumor characteristics are mass margins and mass shape. Respondents considered MRI the better technology to visualize these characteristics. Belief elicitation confirmed this by providing an overall sensitivity of PAM ranging from 58.9 to 85.1 % (mode 75.6 %) and specificity ranging from 52.2 to 77.6 % (mode 66.5 %).

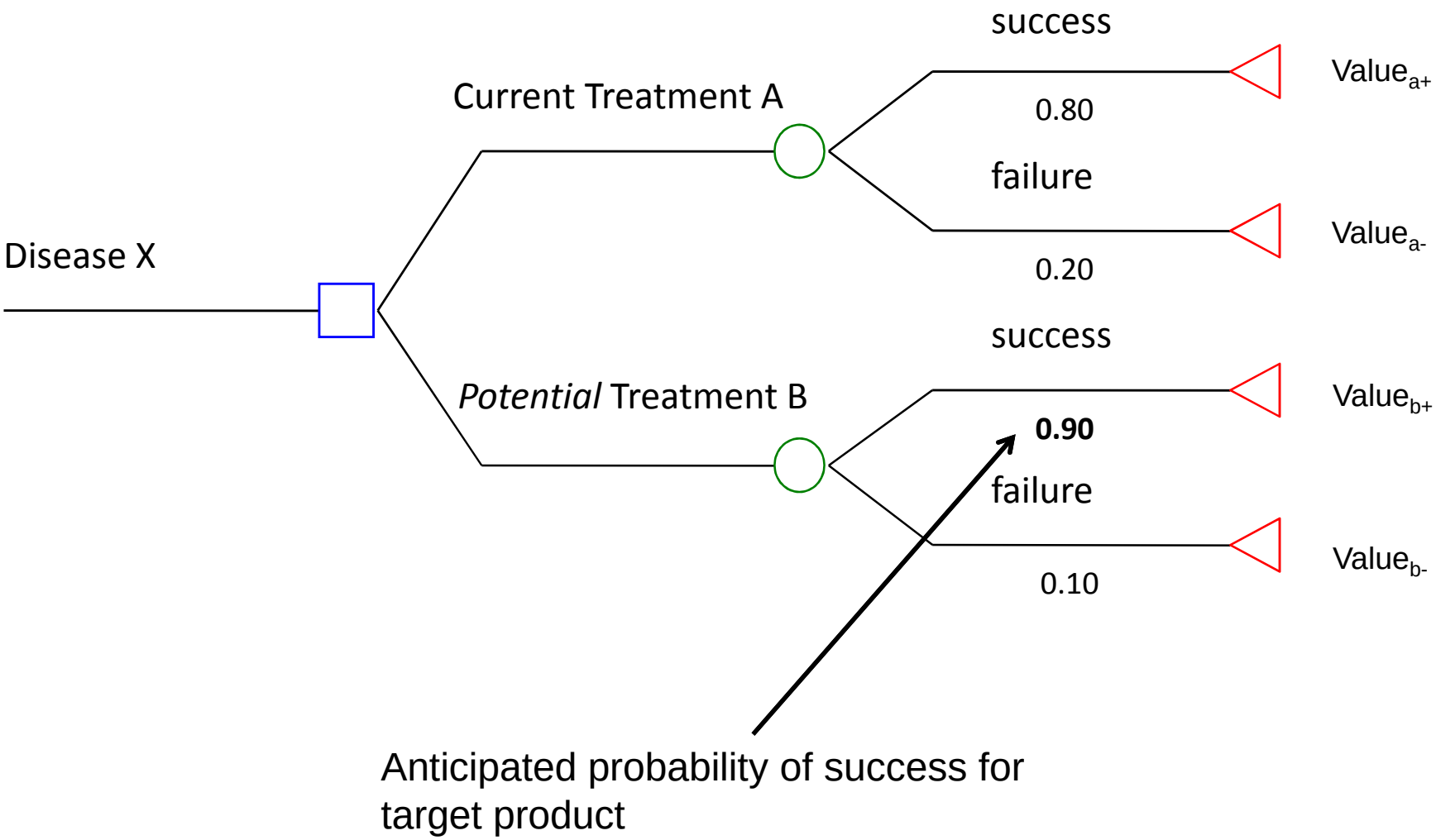
Conclusion Belief elicitation allowed estimates to be obtained for the expected diagnostic performance of PAM, although radiologists expressed difficulties in doing so. Heterogeneity within and between experts reflects this uncertainty and the infancy of PAM. Further clinical trials are required to validate the extent to which this belief–elicitation method is predictive for observed test performance.

Key Points for Decision Makers

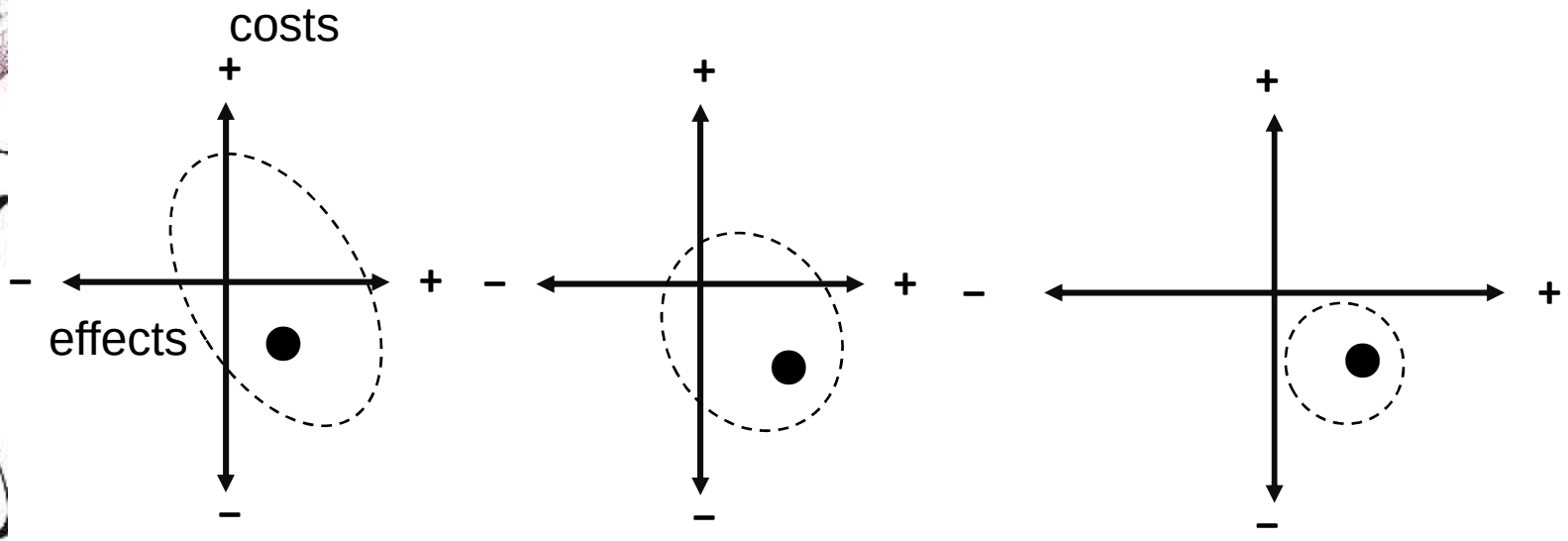
This article presents a new application of belief elicitation to estimate the clinical value of a medical imaging device in an early stage of development.

Belief elicitation in an early stage can identify the

Early Health Economic Modeling: A simple starting point



Analytic option – Bayesian approach



Stadium 1: Few Data

"Cost Effectiveness Gap - analysis"

Stadium 2: Some Data

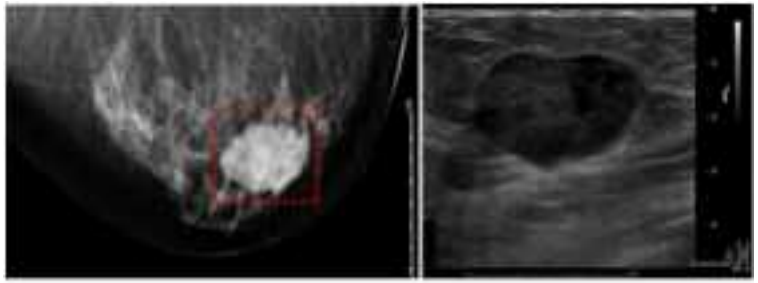
Simple health economic model

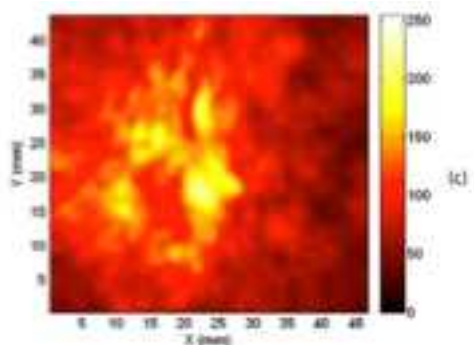
Stadium 3: Rich Data

Full health economic model

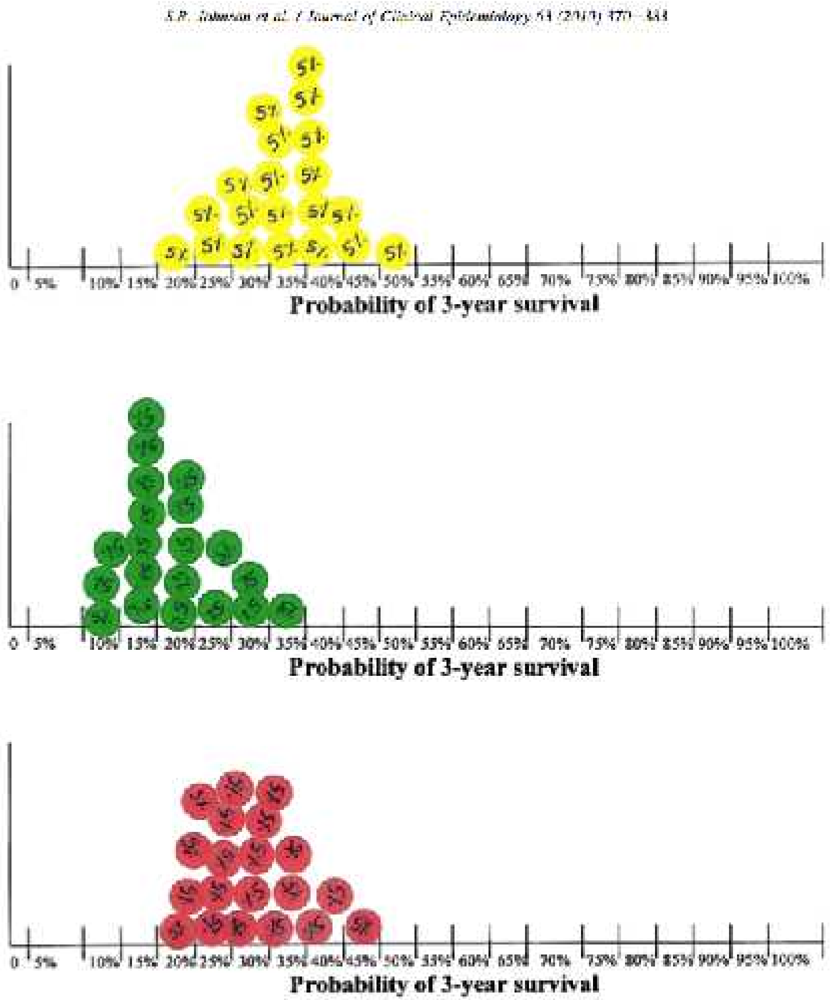
Expert elicitation to estimate uncertain priors

- Case: imaging in breast cancer
- Elicitation of TPR / TNR
- Study design
  - Information on decision context
  - Calibration of radiologists (N=18)
  - Framing of the problem: judging tumor characteristics
- Elicitation experiment
  - TPR and TNR are estimated relative to MRI performance
  - Elicitation of mode and range



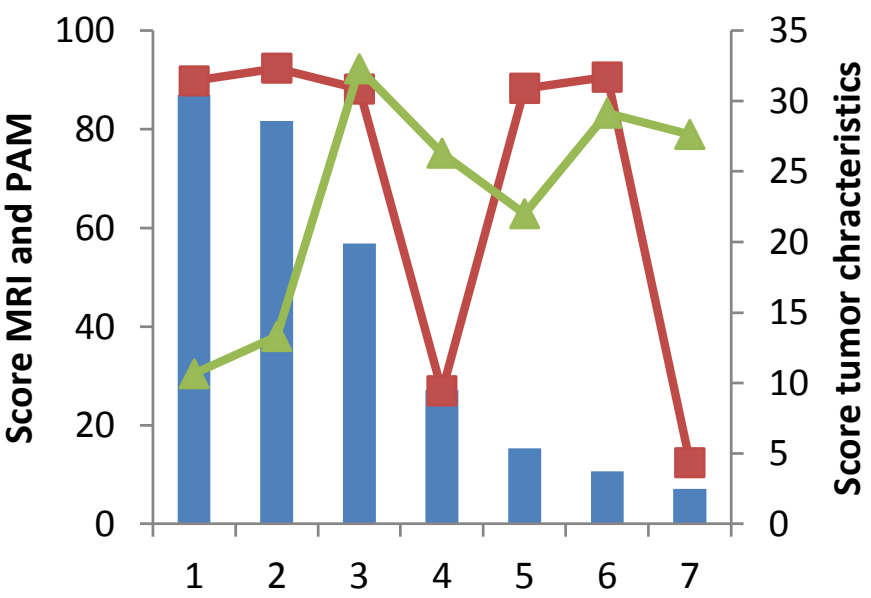


Estimation of priors using expert elicitation



Johnson SR et al: Methods to elicit beliefs for Bayesian priors: a systematic review. Journal of clinical epidemiology. 2010 Apr;63(4):355–69.

Rating of tumor characteristics

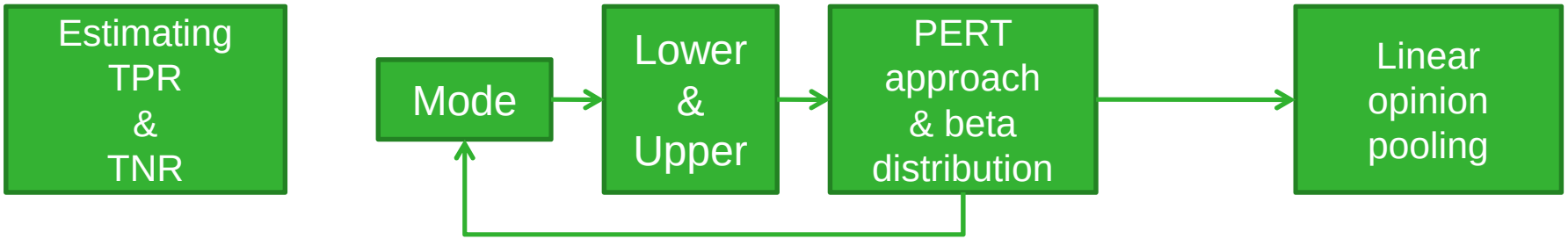


Tumor characteristics
Score MRI
Score PAM

	# of respondents	
	MRI	PAM
1. Mass margins	17	17
2.Mass shape	17	17
3.Vascularization	17	17
4.Mechanical properties	14	13
5.Mass size	17	17
6.Location mass	17	16
7.Oxygen saturation	13	11

Figure 1 Performance and importance tumor characteristics

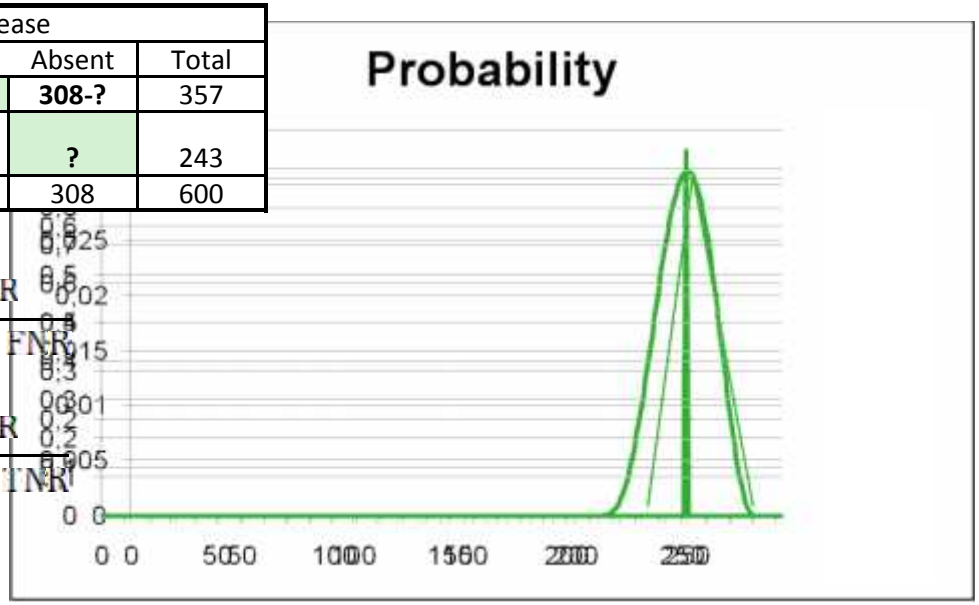
Expert elicitation compared to MRI



		Disease		
		Present	Absent	Total
Test	Positive	**?**	**308-?**	357
	Negative	**292-?**	**?**	243
	Total	292	308	600

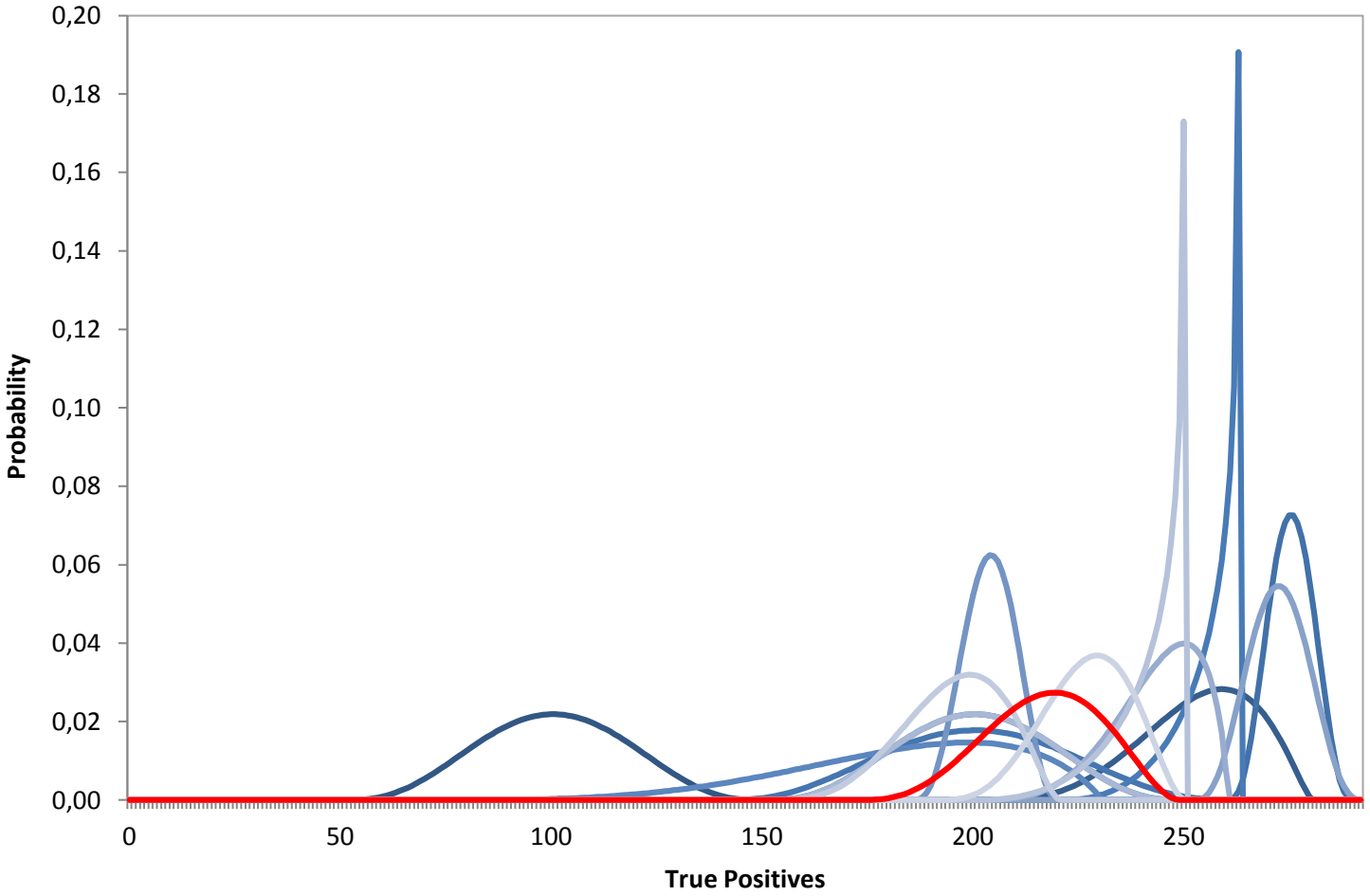
$$\text{sensitivity} = \frac{TPR}{TPR + FNR}$$

$$\text{specificity} = \frac{TNR}{FPR + TNR}$$



Bluemke et al (2004), Bone et al (1996),
Gibbs et al (2004), Nunes et al (2001)

Elicited distributions



Use of probability elicitation in diagnostic imaging

- Considerable heterogeneity between experts
  - Some refused to participate
- Elicited PDF is consistent with overall priority as obtained from rating tumor characteristics
  - Mass margin and shape, consistent with Bi-rads
  - MRI is preferred (sensitivity of 90% compared to 75%)
- Is there a role for probability elicitation?
  - Behavioral approach, i.e. consensus building
  - Decision heuristics, i.e. image processing algorithm
  - Be more specific on case mix, e.g. specific tumor types
  - If more diagnostic information comes available, update expert judgments using clinical vignettes

Lessons learned (1): Similarities and differences

- Distinction between “probability elicitation” and “preference elicitation” i.e. valuing alternatives
- Sample size estimation and sampling of population
  - Facilitated group session (<15) or population based survey
  - Stakeholder groups vs. experts (calibration)
- Measures to avoid (behavioral) biases
  - Informing about context; introduce clinical vignette
  - Value tree: preferential independence
  - Expected range of performance of the alternatives
  - Example or seed-questions
  - Validation of the results with experts

Lessons learned (2): Similarities and differences

- Many differences between methods w/r preference elicitation
  - Direct vs. indirect estimation of preference weights
    - Conjoint analysis methods (DCE, BWS) vs. MCDA
  - Compositional vs. de-compositional methods
  - Compensatory vs. non-compensatory
    - Allowing criteria to be compensated by others
  - Value functions vs. ranking alternatives
    - Aggregate value function
- Guidance being developed
  - ISPOR Taskforces

Lessons learned (3): actual use of expert judgment

- Cases did not actually change R&D choices, but strengthen the ideas and made choices more explicit and transparent
- Most people appreciate facilitated-group discussions for weighing alternatives, i.e. MCDA
- Decision makers seem to value opinions of experts, but are less convinced by quantitative elicitation methods
- Expert opinions is mentioned in pharmaco-economic guidelines (e.g. ZINL) but there is no methodological guidance
-

This presentation

1. Judgment: ‘probability’ or ‘preference’ elicitation
   - **Choice-modeling**: discrete-choice and best-worse scaling
     - CRC screening: patient preferences for alternative screening
     - *ISPOR Taskforces: Bridges, 2011, Johnson, 2013, Hauber, 2015*
   - **MCDA**: weighing, ranking and prioritizing alternatives
     - Photoacoustic imaging; prioritizing further development
     - *ISPOR Taskforces MCDA: www.ispor.org*
   - **Probability elicitation**: elicitation of priors to populate HE models
     - Mammography for breast cancer screening and diagnostics
     - *Johnson et al, 2010 and Butler et al, 2015*
2. Role of expert judgment in personalized healthcare
   - Health systems approach: where to add value?
   - Utility of diagnostic information

2. Expert judgment in personalized healthcare

1. Issues in personalized medicine

- Limited availability of evidence of effectiveness
- The complexity of information provided by tests
- Rapid commercialization of genetic tests
- Apprehension of genomics by physicians
- Uncertainty about their clinical utility

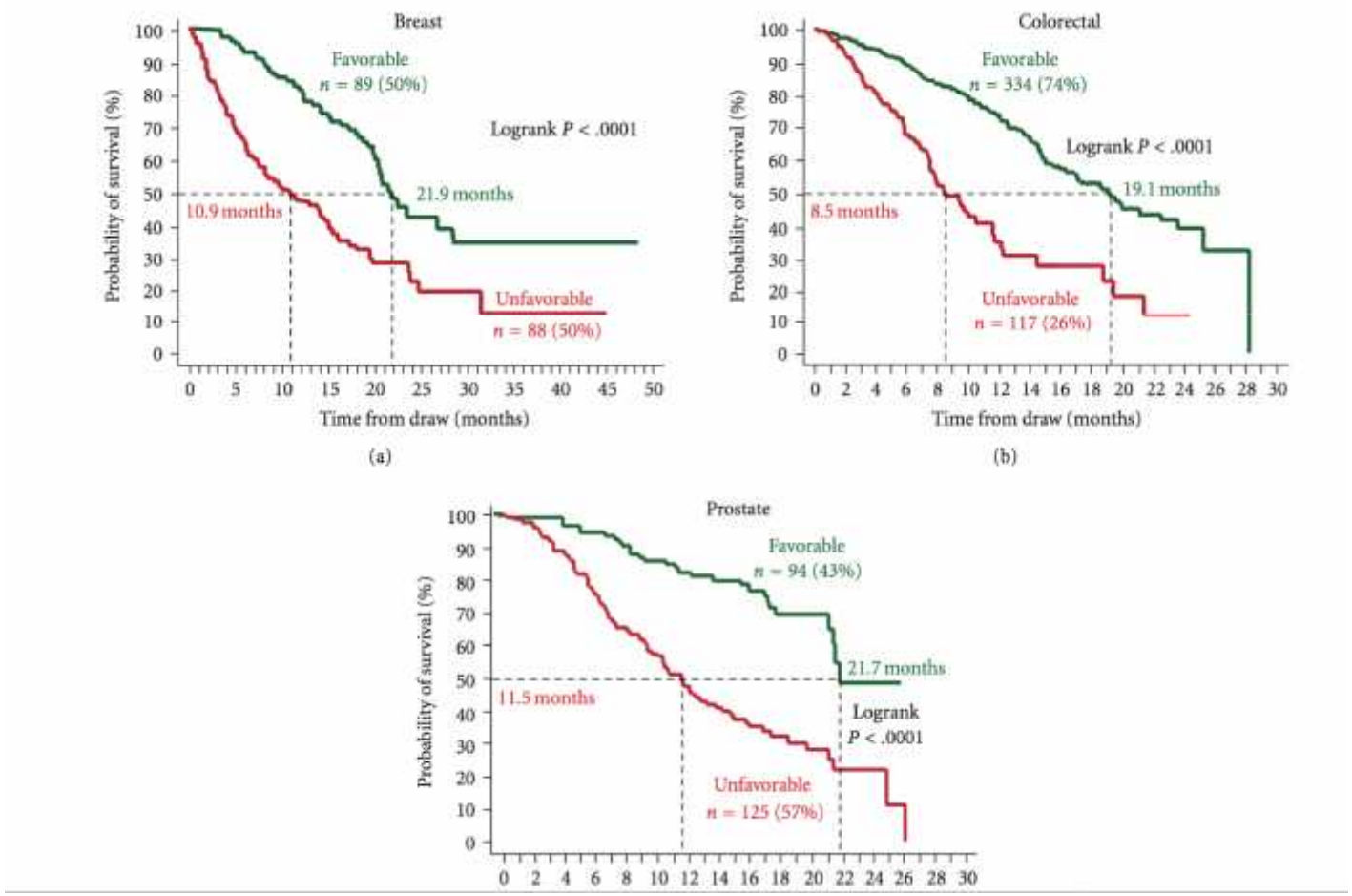
2. Potential for expert elicitation, i.e. diagnostics

- Health systems approach: where to add value?
- Utility of additional diagnostic information
  - Which combination of markers adds value
  - How does it change clinical management?

IJzerman MJ et al; Implementing Comparative Effectiveness Research in Personalized Medicine Applications in Oncology: Current and Future Perspectives. Comparative Effectiveness Research, October; 2015.

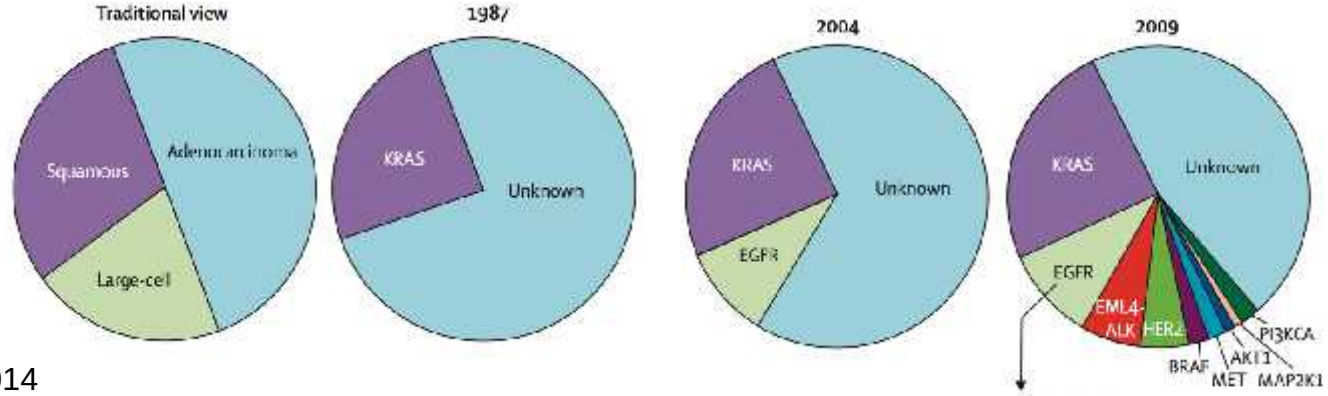
CellSearch© (cleared by FDA)

Prognostic value of CTC counts in different tumors



Miller MC, Doyle GV, Terstappen LWMM. Significance of Circulating Tumor Cells Detected by the CellSearch System in Patients with Metastatic Breast Colorectal and Prostate Cancer. J Oncol. 2010;2010:617421.

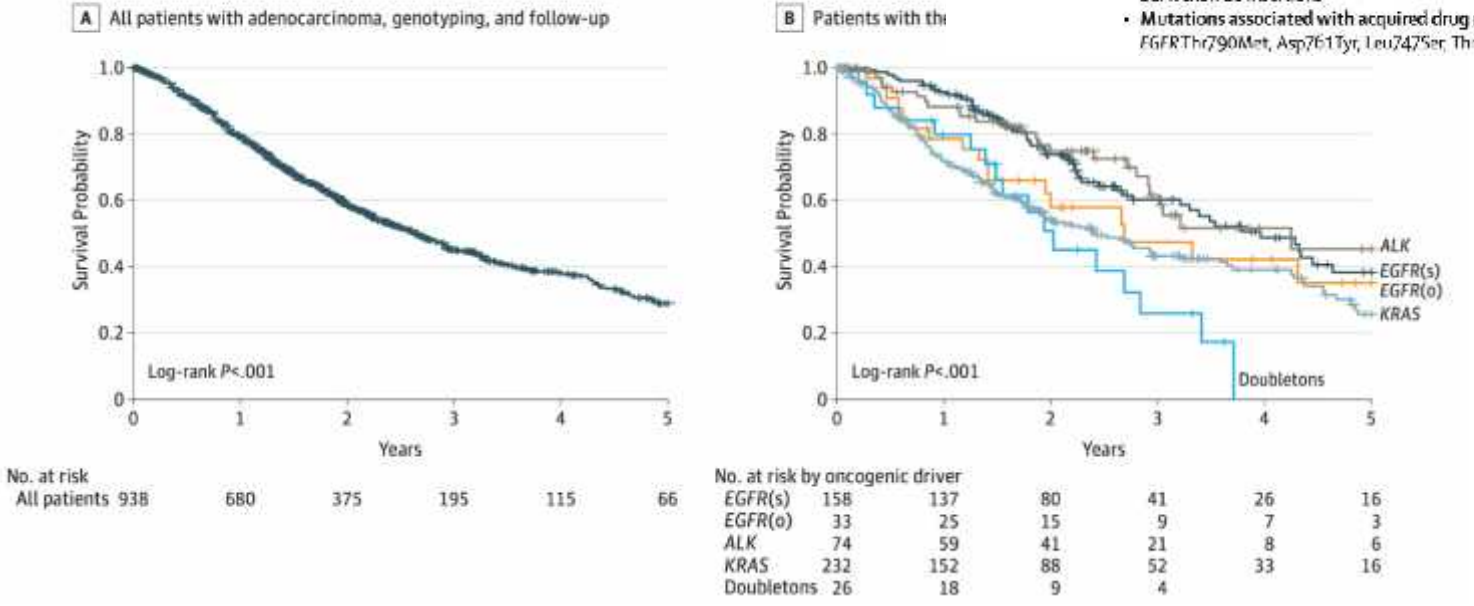
Molecular profiling adds complexity



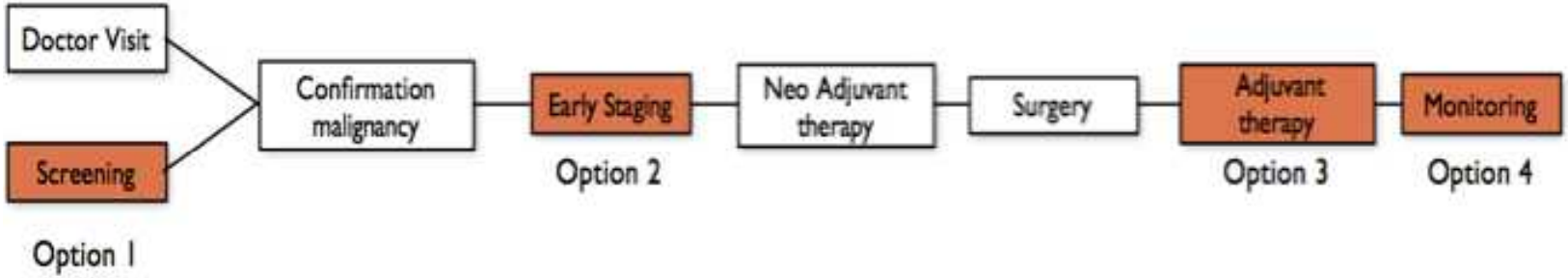
Pao, Lancet 2011
Kris et al, JAMA 2014

- **Mutations associated with drug sensitivity**
  EGFR Gly719X, exon 19 deletion, Leu858Arg, Leu861Gln
- **Mutations associated with primary drug resistance**
  EGFR exon 20 insertions
- **Mutations associated with acquired drug resistance**
  EGFR Thr790Met, Asp761Tyr, Leu747Ser, Thr854Ala

Figure 1. Survival of Patients

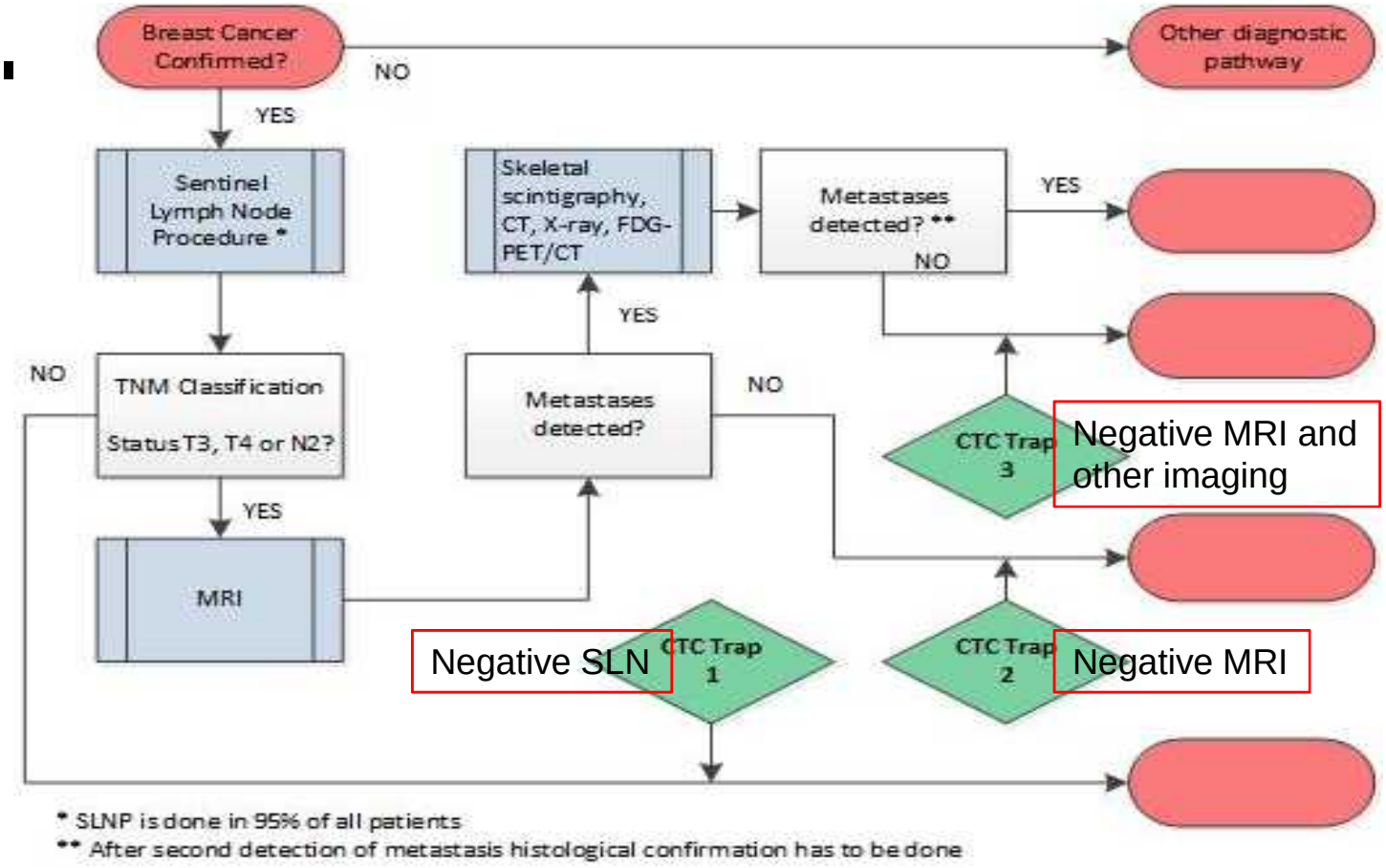


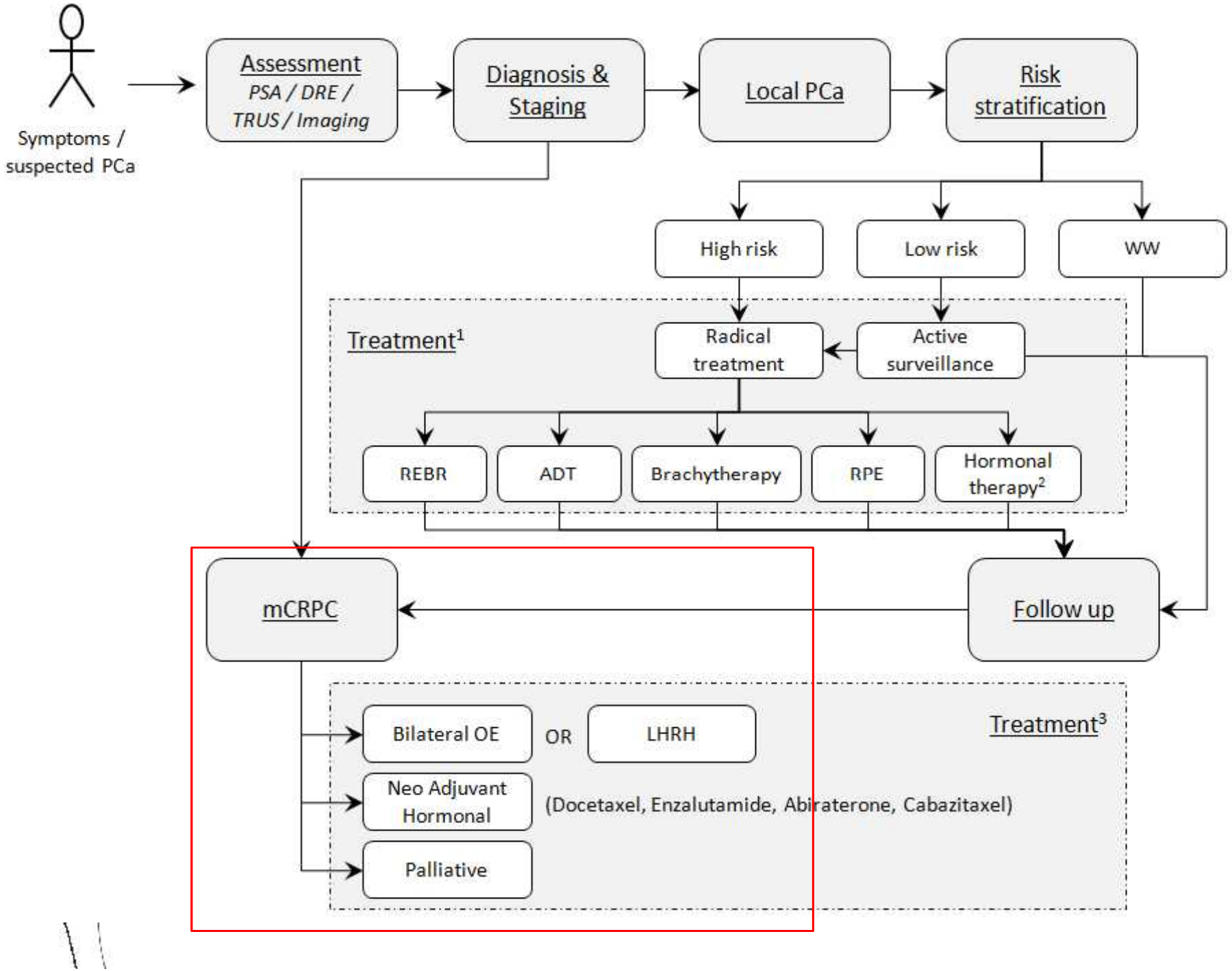
Added value for the health system



963,000 /yr → 12,331 /yr → 656 /yr → 144 /yr

Added value of CTC for staging of breast cancer




Symptoms / suspected PCa
Assessment
PSA / DRE / TRUS / Imaging
Diagnosis & Staging
Local PCa
Risk stratification
High risk
Low risk
WW
Treatment1
Radical treatment
Active surveillance
REBR
ADT
Brachytherapy
RPE
Hormonal therapy2
mCRPC
Follow up
Treatment3
Bilateral OE
OR
LHRH
Neo Adjuvant Hormonal
(Docetaxel, Enzalutamide, Abiraterone, Cabazitaxel)
Palliative

CTC as a response marker: early switching

- Early switching benefits:
  - Early discontinuation and switch to cabazitaxel results in longer PFS and increased quality of life.
  - Early discontinuation may save costs to the health system
  - Early discontinuation of ineffective treatment reduces toxic adverse-events
- 1st and 2nd line therapies (case CTC-Stop trial)

Price per cycle	Docetaxel	Cabazitaxel
UK	£ 534	£ 3696
Netherlands	€ 512 (2 mL)	€ 4646 (1,5 mL)
Canada	$ 599 (16mg/16ml)	$ 5840 (60mg/ml)

“Expert judgment”

And suddenly, you realize how often experts are involved in our research. By asking them to value health services, to prioritize medical technologies and to determine risk tolerance and decision trade-offs.

In the era of personalized healthcare, expert judgment will increasingly add value

- Complexity of chronic diseases, such as dynamic molecular interactions and mutations
- Evidence gaps and uncertainty in clinical data will be more likely
- Big data presents opportunities for risk analysis and clinical management
- Clinical decisions are based on multiple and more detailed information sources; practice guidelines will not be sufficient
- Patient preferences will likely become more important for tailoring health services