Expert Elicitation to Populate Early Health Economic Models of Medical Diagnostic Device Development

Wieke Haakma, MCS1, Laura Bojke, PhD, MSc, BA2, Lotte Steuten, PhD1 and Maarten J. Ijzerman, PhD1
(1) University of Twente, Enschede, the Netherlands, (2) University of York, York, United Kingdom

Introduction

There is an increasing interest to estimate the potential clinical value and likely cost-effectiveness of diagnostic and therapeutic technologies during early development stages to guide further developments. [1,2] Later, early stages of development are typically characterised by large uncertainty and populating health-economic models with empirical data is not always feasible due to limited availability of data. Elicitation of expert opinions is viewed as an appropriate alternative and may serve as the input for early health-economic models. [3]

Figure 1 Flowchart of product development [4]

Objective

In the present study we explore whether expert elicitation is a valid approach to characterise uncertainty regarding the diagnostic performance of photoacoustic imaging in breast cancer. As PAM is still in the translational stages (Figure 1) and the prototype is still in development, there is no clinical data available.

Methods

Different methods have been applied to evaluate medical technologies in early stages of development e.g. Analytic Hierarchy Process (AHP) [4], and expert elicitation. Expert elicitation is intended to link an expression of an expert's beliefs into a statistical format and has been used in Bayesian statistics because of the need to formulate priors.

We have chosen to use expert elicitation as a method to formulate the knowledge and beliefs of experts about the future performance of PAM and to quantify this information into probability distributions.

Table 2 Elicitation factors

Who is an expert?
Mathematical

What to elicit?
Credible interval

Prescribing experts' beliefs
Variable fixed

Probability Distribution Functions
Cumulative Distribution Function

Sample of experts
Twenty radiologists, specialised in the examination of MR images of breasts, from both academic and non academic hospitals in the Netherlands, were invited to participate in this study.

Calibration method
The purpose of calibration is to receive a relative weighting index for each expert. The weight of each individual expert was determined based on clinical background.

Table 2 Calibration factors

Number of experience (years of EXP) | Average number of MRI's examined per week (average EXP) | Examining MRI's in other areas (expert EXP)
--- | --- | ---
K<1 | 1 | No
K=1 | 2 | Yes
K=2 | 3 | Yes/10
K=3 | 4 | No
K=4 | 5 | K=5 | 6 | No/10
K=6 | 7 | Yes
K=7 | 8 | No
K=8 | 9 | Yes
K=9 | 10 | No
K=10 | 11 | Yes
K=11 | 12 | No
K=12 | 13 | Yes
K=13 | 14 | No
K=14 | 15 | Yes
K=15 | 16 | No
K=16 | 17 | Yes
K=17 | 18 | No
K=18 | 19 | Yes
K=19 | 20 | No
K=20 | 21 | Yes
K=21 | 22 | No
K=22 | 23 | Yes
K=23 | 24 | No
K=24 | 25 | Yes
K=25 | 26 | No
K=26 | 27 | Yes
K=27 | 28 | No
K=28 | 29 | Yes
K=29 | 30 | No
K=30 | 31 | Yes
K=31 | 32 | No
K=32 | 33 | Yes
K=33 | 34 | No
K=34 | 35 | Yes
K=35 | 36 | No
K=36 | 37 | Yes
K=37 | 38 | No
K=38 | 39 | Yes
K=39 | 40 | No
K=40 | 41 | Yes
K=41 | 42 | No
K=42 | 43 | Yes
K=43 | 44 | No
K=44 | 45 | Yes
K=45 | 46 | No
K=46 | 47 | Yes
K=47 | 48 | No
K=48 | 49 | Yes
K=49 | 50 | No
K=50 | 51 | Yes
K=51 | 52 | No
K=52 | 53 | Yes
K=53 | 54 | No
K=54 | 55 | Yes
K=55 | 56 | No
K=56 | 57 | Yes
K=57 | 58 | No
K=58 | 59 | Yes
K=59 | 60 | No
K=60 | 61 | Yes
K=61 | 62 | No
K=62 | 63 | Yes
K=63 | 64 | No
K=64 | 65 | Yes
K=65 | 66 | No
K=66 | 67 | Yes
K=67 | 68 | No
K=68 | 69 | Yes
K=69 | 70 | No
K=70 | 71 | Yes
K=71 | 72 | No
K=72 | 73 | Yes
K=73 | 74 | No
K=74 | 75 | Yes
K=75 | 76 | No
K=76 | 77 | Yes
K=77 | 78 | No
K=78 | 79 | Yes
K=79 | 80 | No
K=80 | 81 | Yes
K=81 | 82 | No
K=82 | 83 | Yes
K=83 | 84 | No
K=84 | 85 | Yes
K=85 | 86 | No
K=86 | 87 | Yes
K=87 | 88 | No
K=88 | 89 | Yes
K=89 | 90 | No
K=90 | 91 | Yes
K=91 | 92 | No
K=92 | 93 | Yes
K=93 | 94 | No
K=94 | 95 | Yes
K=95 | 96 | No
K=96 | 97 | Yes
K=97 | 98 | No
K=98 | 99 | Yes
K=99 | 100 | No
K=100

Rating of tumour characteristics
Radiologists are asked to indicate the performance of PAM and MRI for different tumour characteristics used in the examination of images of breasts.

Table 3 Elicitation model data

Positive | Total | Total | Total
--- | --- | --- | ---
Yes | No | Total | Yes | No | Total | Yes | No | Total
Negative | 29 | 234 | 263 | 25 | 25 | 25 | 25 | 25 | 25
Total | 312 | 258 | 600 | 40 | 40 | 40 | 40 | 40 | 40

Elicitation methods

Eliciting distributions
A spreadsheet-based [5] (Excel) exercise was designed to elicit the TPR and TNR. Experts received a face-to-face interview of 30 to 45 minutes in which the similar data regarding PAM was presented to each individual radiologist.

Figure 4 elicitation procedure

Figure 7 shows that there is considerable heterogeneity between radiologists.

Figure 8 Probability distribution of estimations of TPR of 16 radiologists

Conclusion

Experts estimated the mode of the sensitivity and specificity of PAM to be 75.6% and 66.5%, which is lower than MRI (90.5% and 68.5%).

Experts expressed difficulties estimating the performance of PAM based on limited data regarding MRI.

To improve the validity of radiologists’ estimations in this study, it is desirable to elicit priors for specific tumor types, since radiologists indicated to base their estimations on an aggregate expectation about how MRI will visualize the various tumor types.

Further clinical trials should be commissioned to indicate whether these results are valid and expert elicitation could be used in early technology assessment. Before that, the use of the elicited priors in economic models requires careful consideration.

 References