PRIORITIZATION AND WEIGHTING OF PATIENT-RELEVANT ENDPOINTS AS PART OF IQWIG'S EFFICIENCY FRONTIER METHOD IN GERMANY

Axel C. Mühlbacher[1,2], John F.P. Bridges[3], Susanne Bethge[1,2], Ch.-M. Dintsios[4], A. Schwalm[4], M. Nübling[5]

[1]Duke Clinical Research Institute, The Fuqua Business School, Duke University, USA; [2]Institute Health Economics and Health Care Management, University of Applied Sciences Neubrandenburg, Germany; [3]Johns Hopkins Bloomberg School of Public Health, Baltimore, [4] Institute for Quality and Efficiency in Health Care, Germany, [5] Gesellschaft für empirische Beratung mbH, Germany

Background, Objectives and Results

The methods concept for health economic evaluation envisages the application of an efficiency frontier to estimate the relationship between the benefits and costs of a medical intervention compared to a relevant alternative within a therapeutic area. The benefit is assessed by the Institute for Quality and Efficiency in Health Care (IQWiG) by means of patient-relevant endpoints (PREs). According to Social Code Book V (SGB V), the endpoints primarily used are mortality, morbidity and health-related quality of life. IQWiG allows the inclusion of various endpoints and thus the analysis of different efficiency frontiers. If decision-making is based on several efficiency frontiers, it remains open how the results were interpreted by health policy decision makers.

Objective: Focus 1: It will be outlined how the conjoint / discrete choice analysis can be used to identify PREs. Moreover, it will be shown how the method can be applied in benefit assessments, using antiviral hepatitis C therapy as an example. It will also be tested whether this method is applicable for weighting and prioritization across multiple endpoints. In this context, simultaneously the practicability and applicability of the approach (validity and plausibility of results) will be tested and displayed. Focus 2: It will be demonstrated how, by means of the conjoint / discrete choice analysis, an approximative cardinalization of PREs can be presented. Focus 3: Finally the validity and reliability of discrete choice models in the assessment of multiple endpoints will be discussed.

Research questions: This methodological approach allows the discussion of the following research questions: identification and prioritization of PREs; approximative cardinalization: testing of the linearity assumption; patient populations: testing of the heterogeneity assumption; patient preferences versus expert judgments.

In addition to the objectives of the pilot study, for the comprehensive presentation of the application options of the discrete choice method, two additional research questions were examined. Firstly, for prioritization: the maximum willingness to accept the risk of adverse effects (maximum acceptable risk). Secondly, for determination of an endpoint-based value: the comparison of therapy alternatives.

Methods: The qualitative phase of the study comprised focus groups as well as a pre-test questionnaire. Within the quantitative phase, a survey of patients and experts was conducted with paper-and-pencil and/or online questionnaires. A main-effects orthogonal design was used based on an orthogonal array (6^7) with a total of 72 choice sets. Each expert assessed 72 choice sets. For the patients the design was divided into 4 blocks with 18 choice sets each.

Frequencies and statistical characteristics of distributions were used, as well as bivariate (ANOVA) and regression analyses, and logit and probit models (SPSS 18 und STATA 11). Within the model calculation effect codes where used. The assessment of model quality was performed by means of Akaike's Information Criteria (AIC) and Bayesian Information Criteria (BIC), the likelihood ratio test, and the percentages of choice sets correctly predicted.

Results: The main survey was conducted in September and October 2010 and included 326 patients and 21 experts. (Tab. 3) Patients and experts weighted endpoints in the same order but with different strengths. The endpoint "sustained virological response" (SVR) was weighted the highest both by patients and by experts. (Tab. 1-2) Decisions on the balancing of benefits and harms differed considerably between experts and patients. (Fig. 3)

Using the examples of 3 therapy alternatives, a potential path to determine an endpoint-based benefit value is outlined. (Tab. 4-6)

State of the Art and Critique of the Efficiency Frontier Method

Identification, weighting and prioritization of multiple endpoints: the health economic evaluation usually includes outcome and treatment situations (e.g. patient subgroups, therapeutic areas) for which an additional benefit (or less harm) compared to other treatment options was shown in the benefit assessment (IQWiG 2009a). When identifying PREs it needs to be considered that patients mostly do not define direct clinical endpoints as goal of treatment, but rather an improvement in "health" or an increased quality of life resulting from therapy (Breyer 2010).

In the decision-making process the relationship between the benefit and the risk of adverse effects may play an essential role. In this context it would be meaningful to analyze the willingness of patients to accept the risk of adverse effects (risk preferences of patients; maximum acceptable risk). The patient benefit can only be assessed with involvement of the affected patients themselves. As a one-dimensional indicator to explain choices about decisions, patient preferences represent the extent to which a therapy alternative should be favored from the patient perspective. The perceived or expected benefit (or harm) from the perspective of the patient is thereby the basis for patient preferences and consequently one of the explaining factors for a patient's choice of action. For this reason, on the basis of the known preferences of an individual or a patient population, conclusions can be drawn about the benefit of a treatment alternative.

Approximative cardinalization: When presenting efficiency frontiers it needs to be considered that the value of a clinical endpoint (for the patient) should be entered on the ordinate for construction of the frontier, not the actual clinical result, as has so far been common practice. Cardinal measurability must thus be assumed not only for the difference in results but also for the difference in values between the endpoint levels. In the context of an allocation decision founded on health economics and welfare economics, it is important that the value differences of the endpoints are used for the assessment. Drummond and Rutten (2008) also state that ultimately the whole informational content of the efficiency frontier depends on whether the values displayed on the y axis are available in cardinal units.

Subgroup analyses (heterogeneity): In the assessment of endpoints different weighting factors for patient benefit may arise due to factors such as gender, age, disease status, disease stage, concomitant illnesses, and risk factors. It is unclear how these subgroup differences can be determined and presented, how such differences can be handled, and what implications arise from them for the decision-making process (Sculpher, Claxton 2010). Heterogeneous patient preferences in different populations may have an essential impact on the patient benefit.

Expert judgments versus patient preferences: Every conceivable perspective of stakeholders should be taken within the framework of an assessment procedure, i.e. depending on the decision maker, the perspective of citizens, the social security community, the patient, the insured person, or the expert. In this context it can be examined to what extent the opinions of the various groups differ and what impact this has on the assessment of the benefit.

Results

Patients:

Tab. 1: Coefficients and odds ratios of logit model patients

Attribut	Coeff.	Odds ratio	se coeff	sig	95%CI low	95%CI up	95%CI oneway
(1) duration of treatment	0,2502	1,284282	0,234200	< 0.001	0,2043	0,2962	0,0459
(2) frequency of injecting interferon	0,2966	1,345277	0,233700	< 0.001	0,251	0,3424	0,0456
(3) duration of flue like symptoms after injection	0,1052	1,110933	0,232300	< 0.001	0,06	0,1507	0,0452
(4) probability of getting gastrointestinal symptoms	0,1233	1,131224	0,234300	< 0.001	0,078	0,169	0,0453
(5) probability of getting psychiatric symptoms	0,1857	1,204061	0,232700	< 0.001	0,1398	0,2317	0,0459
(6) probability of getting skin problems or Alopecia	0,1054	1,111155	0,261100	< 0.001	0,0599	0,1511	0,0455
(7) probability of sustained virological response 6 month after treatment	0,8041	2,234684	0,261000	< 0.001	0,75295	0,8553	0,05115

Random-effects logistic regression, Number of obs = 5252, Number of groups = 309, LR chi2(7) = 1563.26,
Log likelihood = -2852.7476, Prob > chi2 = 0.0000

Experts:

Tab. 2: Coefficients and odds ratios of logit model experts

Attribut	Coeff.	Odds ratio	se coeff	sig	95%CI low	95%CI up	95%CI oneway
(1) duration of treatment	0,7918	2,207451	0,069329	< 0.001	0,6560	0,9277	0,1358817
(2) frequency of injecting interferon	0,4053	1,499740	0,056374	0,0000	0,2948	0,5158	0,1104905
(3) duration of flue like symptoms after injection	0,0786	1,081725	0,056094	0,1610	-0,0314	0,1885	0,1099418
(4) probability of getting gastrointestinal symptoms	0,1620	1,175813	0,058546	< 0.01	0,0472	0,2767	0,1147476
(5) probability of getting psychiatric symptoms	0,2702	1,310261	0,059416	< 0.001	0,1538	0,3867	0,1164531
(6) probability of getting skin problems or Alopecia	0,0622	1,064199	0,058534	0,2880	-0,0525	0,1769	0,1147253
(7) probability of sustained virological response 6 month after treatment	1,7362	5,675621	0,086153	< 0.001	1,5673	1,9050	0,168857

Random-effects logistic regression, Number of obs = 1512, Number of groups = 21, LR chi2(7) = 1076.62,
Log likelihood = -509.20122, Prob > chi2 = 0.0000

Fig. 1: Approximative cardinalization, Patients

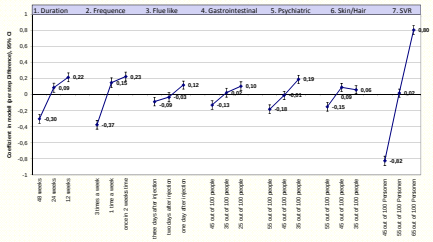
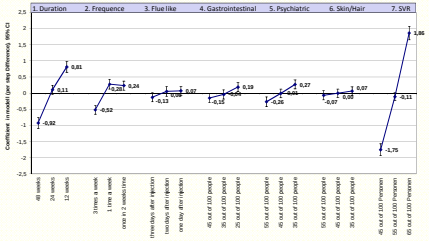


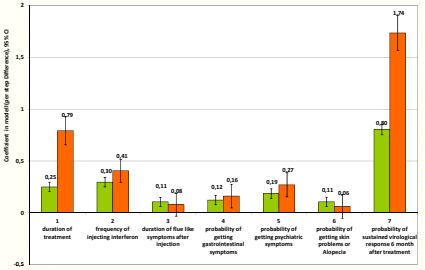
Fig. 2: Approximative cardinalization, Experts



Tab. 3: Sociodemographics of Study Samples

Patients		Experts	
Gender	**N**	**Gender**	**N**
valid	325	*valid*	21
Female	208	Female	2
Male	117	Male	19
Age	**N**	**Age**	**N**
valid	319	*valid*	21
minimum	16	minimum	36
maximum	88	maximum	65
mean	58,89	mean	47,95
missing	7	missing	0
Current treatment status	**N**	**Possition**	**N**
valid	325	*valid*	21
never started a treatment	92	intern	1
currently under treatment	25	Attending	7
finished treatment	130	chief of medicine	3
discontinuation of treatment	78	Other (eg. primary care physician)	10
missing	1	missing	0

Fig. 3: Comparison of Patients and Experts Coefficients



Further Results

Determination an endpoint-based value – The connections between evidence-based medicine and health economics: Using 3 possible treatment alternatives as examples, a path to determine an endpoint-based value is outlined. The aim is the calculation of a cardinally scaled benefit dimension. The presented approach for determining an endpoint-based value is directly based on the principles of evidence-based medicine. The model is variable and can be adapted to future innovations (as long as the assessment is to be based on the same endpoints).

Current treatment standards served as the basis of comparison for the present study. Using the weightings of endpoints determined in the present study results in the following 3 endpoint-based values (in each case according to the type of therapy used). (Tab. 4-6)

Tab. 4: Endpoint-based value assessment therapy A: Peginterferon (high dose) + Ribavirin

Endpoint	E_n	E_{Min}	E_{Max}	Z_{nm}	G_{nm}	T_{nm}
1) duration of treatment	48	48	48	0	0,2502	0
2) frequency of injecting interferon	1	1	3	0	0,2966	0
3) duration of flue like symptoms after injection	2	2	2	0	0,1502	0
4) probability of getting gastrointestinal symptoms	27,75	22,25	27,75	100	0,1233	-12,33
5) probability of getting psychiatric symptoms	30,75	29,75	32,50	36,36	0,1857	-6,75
6) probability of getting skin problems or Alopecia	32,67	26,67	32,67	100	0,105	-10,5
7) probability of sustained virological response 6 month after treatment	54	47	54	100	0,8041	80,41
Endpoint-based value assessment therapy A						**50,83**

Tab. 5: Endpoint-based value assessment therapy B: PegInterferon (low dose) + Ribivarin

Endpoint	E_n	E_{Min}	E_{Max}	Z_{nm}	G_{nm}	T_{nm}
1) duration of treatment	48	48	48	0	0,2502	0
2) frequency of injecting interferon	1	1	3	0	0,2966	0
3) duration of flue like symptoms after injection	2	2	2	0	0,1502	0
4) probability of getting gastrointestinal symptoms	23,75	22,25	27,75	27,27	0,1233	-3,36
5) probability of getting psychiatric symptoms	29,75	29,75	32,50	0	0,1857	0
6) probability of getting skin problems or Alopecia	30,17	26,67	32,67	58,33	0,105	-6,12
7) probability of sustained virological response 6 month after treatment	47	47	54	0	0,8041	0
Endpoint-based value assessment therapy B						**-9,48**

Tab. 6: Endpoint-based value assessment therapy C: Interferon + Ribivarin

Endpoint	E_n	E_{Min}	E_{Max}	Z_{nm}	G_{nm}	T_{nm}
1) duration of treatment	48	48	48	0	0,2502	0
2) frequency of injecting interferon	3	1	3	100	0,2966	-29,66
3) duration of flue like symptoms after injection	2	2	2	0	0,1502	0
4) probability of getting gastrointestinal symptoms	22,25	22,25	27,75	0	0,1233	0
5) probability of getting psychiatric symptoms	32,5	29,75	32,50	100	0,1857	-18,57
6) probability of getting skin problems or Alopecia	26,67	26,67	32,67	0	0,105	0
7) probability of sustained virological response 6 month after treatment	47	47	54	0	0,8041	0
Endpoint-based value assessment therapy C						**-48,23**

Conclusion

The approach to a benefit assessment presented here is based on the principles of evidence-based medicine. In addition, its derivation from utility theories has the advantage of a close connection to microeconomic research, so that conclusions about welfare economics can also be drawn from the model's results. The potential for application was shown in the following areas:

Identification, weighting and prioritization of multiple endpoints: The discrete choice experiment calculates the weighting factors of the individual patient-relevant endpoints. The increasing necessity of health economic evaluations with particular consideration of patient-relevant endpoints is particularly evident against the background of the latest developments in antiviral therapy of chronic hepatitis C (completion of Phase III of the Vertex RCT) and the potential market of new drugs, which have considerably improved SVR rates (and a concurrent reduction in treatment duration), as well as better response rates in interferon- and ribavirin-resistant patients (Hay 2010; O'Leary, Davis 2010; Pockros 2010). (Tab.1 -2)

Approximative cardinalization of clinical effect sizes: For the following endpoints, a linear increase in levels across the levels used can be shown: "probability of SVR 6 months after end of treatment", "duration of flu-like symptoms after injection", "probability of gastrointestinal symptoms", and "probability of psychiatric symptoms". However, for the endpoint "probability of skin symptoms and/or alopecia" a linear course cannot be assumed. The decisive factor for the benefit assessment is not the difference of clinical effect sizes, but the value difference of these endpoints. The results presented allow the testing of this assumption, an assumption that is necessary to draw the efficiency frontier. (Fig. 1-2)

Subgroup analyses (heterogeneity): In patients with severe HCV infections it could be demonstrated that weighting is performed in favor of the expected therapy success and that toxicity becomes less relevant. It could be shown that, in particular with regard to the 3 attributes with the highest weighting, subgroup effects existed. This applied to the following model variables: gender, age, marital status, education, occupation, net income, stage of fibrosis, genotype, year of first diagnosis, and prior treatment with antiviral drugs.

Expert judgments versus patient preferences: In this study the order of weightings was largely congruent between patients and experts. However, for the patient-relevant endpoints "duration of antiviral treatment" "probability of psychiatric symptoms" "frequency of interferon injections" minor deviations regarding the order and strength of weighting were shown. For example, experts judged adverse effects to have a far smaller impact on treatment or the choice of treatment than patients did. (Fig. 3)

Further application options:

Determination of an endpoint-based value: Assuming a linear function of benefit, the clinical results of different treatment alternatives can be aggregated. The endpoint-based value can be entered on the benefit axis and an efficiency frontier across endpoints can thus be displayed. (Tab. 4-6)

Patient perspective on the benefit-harm relationship: Using the example of a benefit-harm relationship of an intervention, it was shown that prioritization of attributes can also be performed by means of the maximum acceptable risk of adverse effects. In the analysis of the benefit-harm relationship from the patient perspective, it was shown patients are prepared to accept more adverse effects in exchange for a greater change of treatment success.

Acknowledgement: This Study was founded by the IQWiG. References are available upon request by the authors.