IMPACT HTA

Improved methods and actionable tools for enhancing HTA

Work Package 7 Deliverable 3 (D7.3): Testing the IMPACT-HTA Value Framework in collaboration with HTA agencies: Case studies on Non-Small Cell Lung Cancer and Spinal Muscular Atrophy

EXECUTIVE SUMMARY

Aris Angelis (LSHTM, LSE), Monica Oliveira (IST), Teresa Rodrigues (IST), Liliana Freita (IST), Carlos Bana e Costa (IST) and Panos Kanavos (LSE)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 779312. The results presented reflect the author's views and not those of the European Commission.

Executive summary

Earlier research in Work Package 7 (WP7) led to the development of the IMPACT-HTA Value Framework, a quantitative decision analytic methodological approach, to help HTA agencies and evaluation committees using more structured and consistent processes in the evaluation of new medicines. Here we describe the adaptation of the IMPACT-HTA Value Framework with two HTA agency partners, the Swedish TLV and the Belgian INAMI, and its empirical testing with simulated HTA evaluation committees for two cases studies, one on Non-Small Cell Lung Cancer (NSCLC), and one on Spinal Muscular Atrophy (SMA).

Overall, four main methodological stages comprised a collaborative value modelling (CVM) approach, using a combination of preference elicitation means, including web-Delphi processes and decision conferences as part of a multi-criteria decision analysis (MCDA) methodology. Eventually, the framework was tested in practice with the case studies, which were in alignment with the ISPOR MCDA Good Practice Guidelines, following a two-step interactive value modelling process (an online web-questionnaire survey followed by a decision conference), engaging a group of experts from each HTA agency that simulated evaluation committees.

In the web-Delphi process engaging INAMI and TLV stakeholders, a preliminary generic list of 20 value aspects was supplemented with 3 additional aspects for INAMI, and 4 additional aspects for TLV, giving a total of 23 and 24 value aspects respectively, acting as the final HTA agency-specific lists of generic value aspects. In the first decision context (i.e. indication expansion in severe disease with a number of treatments available for INAMI, malignant neoplasm metastases for TLV), 12 value aspects were classified as “Essential”, 3 as “Influential”, and 8 as “Complementary” for INAMI; no value aspect was judged to be “Irrelevant”. For TLV, 5 value aspects were classified as “Essential”, 7 as “Influential”, 10 as “Complementary” and 2 as “Irrelevant” for TLV. In the second decision context (i.e. life threatening orphan disease with few symptomatic treatments and entry of new advanced therapy for INAMI, or motor neuron diseases and related disorders for TLV), 11 value aspects were classified as “Essential”, 6 as “Influential”, and 6 as “Complementary” for INAMI; again, no value aspect was judged to be “Irrelevant”. For TLV, 8 value aspects were classified as “Essential”, 7 as “Influential”, 7 as “Complementary” and 2 as “Irrelevant” for TLV.

The most substantial (proportional) differences between the decision contexts within the same HTA agency were observed for the value aspects classified as “Influential” with INAMI (n = 3 vs 6),

followed by the value aspects classified as “Essential” with TLV (n = 5 vs 8). In terms of the “Complementary” classification, TLV judged a larger number of such value aspects in both contexts (n = 10; n = 7) compared to INAMI (n = 8; n = 6). After combining the “Essential” and “Influential” value aspects together in a “high relevance” category and the “Complementary” and “Irrelevant” value aspects together in a “low relevance” category, a total of 15 “high relevance” value aspects and 8 “low relevance” value aspects were generated from INAMI in the first decision context, vs 12 and 12 from TLV respectively. Similarly, a total of 17 “high relevance” value aspects and 6 “low relevance” value aspects were generated from INAMI in the second decision context, vs 15 and 9 from TLV, indicating a larger number of “high relevance” value aspects with INAMI compared to TLV, and a larger number of “low relevance” value aspects with TLV compared to INAMI, in both decision contexts.

The two-step interactive process for the NSCLC case study initially involved a web-questionnaire to collect individual value judgments about the performance of three immunotherapies (nivolumab, pembrolizumab and atezolizumab), relative to a chemotherapy (docetaxel), on five clinical value aspects. Nine INAMI and 5 TLV participants expressed separately their individual value judgements, with group majority judgements derived from them, based on which the prototype multicriteria value model was built. Following the first web-questionnaire step, an online decision conference took place (10 participants with INAMI, 5 participants with TLV) to develop a shared multicriteria model and discuss its outputs, enhancing the prototype model. Compared to the prototype model, treatments’ ranking based on the final results following the INAMI decision conference was changed, with Atezolizumab becoming second and Pembrolizumab dropping to third. Although differences also existed in the group’s value judgements following the TLV decision conference, any emerging changes in value scores and weights did not impact the treatment’s ranking. Overall, differences existed in the final rankings between the two agencies, with Atezolizumab being ranked second followed by Pembrolizumab third in the case of INAMI, and vice versa for the case of TLV; with both agencies, Nivolumab ranked first, with docetaxel being last. In terms of relative weights, for the case of INAMI more than 1/3 of the model’s total weight corresponded to the Adverse events profile (36.3%), with about 1/4 of the model’s total weight corresponding to Duration of treatment effect(s) (22.7%), and then an equal weight for Impact of mortality, Impact on morbidity and Tolerability to patients (13.6%). For the case of TLV, just over 1/3 of the model’s weight corresponded to Impact on mortality (34.4%), with 1/4 corresponding to Duration of treatment effect(s) (25%), just over 1/5

corresponding to Impact on morbidity (21.8%), 1/8 to Adverse events profile (12.5%) and less than 1/10 to Tolerability to patients (6.2%).

Similarly, the two-step interactive process for the SMA case study initially involved a web-questionnaire to collect individual value judgments about the performance of the treatments of interest (Zolgensma, Spinraza, and Evrysdi), relative to Best Supportive Care (BSC), on four or five clinical value aspects. Eight INAMI and 6 TLV participants expressed separately their individual value judgements, with group majority judgements being derived from them, based on which the prototype multicriteria value model 1 was built. Following the first web-questionnaire step, a virtual decision conference took place (8 participants with INAMI, 5 participants with TLV) to develop a shared multicriteria model and discuss its outputs, enhancing again the preliminary prototype model. Although differences existed in the group's value judgements following the INAMI decision conference, any emerging changes in value scores and weights did not lead to a difference in the treatment's ranking. For the case of TLV, compared to the prototype model, treatments' ranking based on the final results following the decision conference was changed, with Evrysdi becoming second and Spinraza dropping to third. Overall, differences existed in the final (clinical) rankings between the two agencies, with Evrysdi being ranked first followed by Zolgensma in the case of INAMI, and vice versa for the case of TLV; with both agencies, Spinraza ranked third, followed by BSC. In terms of relative weights, for the case of INAMI more than 1/3 of the model's total weight corresponded to the Impact on mortality (35.7%), with an identical weight equal to 1/4 of the model's total corresponding to Tolerability to patients (25%) and Health Related Quality of Life (25%), and then Impact on morbidity (14.3%). For the case of TLV, just under 1/3 of the model's weight corresponded to Health Related Quality of Life (31.7%), and Impact on mortality (29.3%), with just over 1/5 corresponding to Impact on morbidity (22%), about 1/8 to Tolerability to patients (12.2%) and almost 1/20 to Tolerability to patients (4.9%). In terms of the SMA non-clinical value aspects in model 2, for the case of INAMI instead of Evrysdi which was the top-ranked option in terms of clinical value aspects, Zolgensma was the top-ranked option in terms of non-clinical value aspects, followed by Evrysdi and then by Spinraza. For the case of TLV, Zolgensma was also the top-ranked option in terms of non-clinical value aspects, having the same ranking with clinical value aspects.

The results indicate that what matters in the evaluation of new medicines for HTA agencies is not limited to their clinical benefits in terms of mortality and morbidity but represents a more comprehensive concept, which could be captured by several clinical and non-clinical value aspects.

This value concept becomes more complex given that value aspects and their degrees of relevance also vary between decision contexts, both in terms of evaluation within the same agencies or committees, but also between them. Therefore, HTA institutions would benefit from methodological tools than can be flexible enough to serve the individual needs of different HTA agencies but on the same time they should be characterized by some common value benchmarks in order to draw comparisons between evaluations.

The IMPACT-HTA value framework is characterized by a common but flexible value frame which can be adapted according to individual HTA agency needs. It is able to capture what matters to all relevant stakeholders and also by how much, given the ability to incorporate and elicit value trade-offs as part of an engaging socio-technical approach. Future research could focus on further empirical applications of the IMPACT-HTA value framework, in collaboration with HTA agencies and for different decision contexts, in order to better understand its potential benefits and limitations.

The material included in the full deliverable D7.3 is currently being considered for publication in the peer review literature. As the review process is ongoing, the full D7.3 will be made available on the IMPACT-HTA website (www.impact-hta.eu) in January 2023, once the peer review process has been completed and the papers associated with D7.3 have been published.