PharmacoEconomics

Guidance for the harmonisation and improvement of economic evaluations of personalised medicine

Heleen Vellekoop*[1] MSc, Simone Huygens[1] PhD, Matthijs Versteegh[1] PhD, László Szilberhorn[2] PhD, Tamás Zelei[2] PhD, Balázs Nagy[2] PhD, Rositsa Koleva-Kolarova[3] PhD, Apostolos Tsiachristas[3] PhD, Prof Sarah Wordsworth[3] PhD, Prof Maureen Rutten-van Mölken[1,4] PhD, on behalf of the HEcoPerMed consortium

*Correspondence: vellekoop@imta.eur.nl

[1] Institute for Medical Technology Assessment, Erasmus University Rotterdam

[2] Syreon Research Institute

[3] Health Economics Research Centre, University of Oxford

[4] Erasmus School of Health Policy & Management, Erasmus University Rotterdam

Online Resource 4 – Summarised interview responses

Summarised responses for health economic modellers

INTERVIEWEE	PARAPHRASED RESPONSE
Effectiveness data from non-randomised (controlled) studies	
IC1	Matching (such as propensity score matching) is complicated in increasingly small sample sizes since predictors are not easily identified due to a lack of statistical power
IC2	Small sample sizes increase the need for international collaboration on data sharing, such as registries, but the local governance of these registries often does not allow the pooling of international data
IC3	Many different methods are used to estimate the missing control arms of studies, but often only the preferred method is reported, which may give insufficient attention to the structural uncertainty in the selection of indirect comparison method.
IC4	While RCTs may be complicated, they are not impossible to conduct, rather they are difficult to conduct. PM, in no way, precludes conducting RCTs and following this narrative does not do justice to the true possibilities out there for conducting complicated trials. The costs of not doing these trials may be large. The main issues in PM is small sample sizes and uncontrolled data and the classic issues that arise from that, such as the unknown natural course of the disease, which is often unknown for specific new genetic mutations that are identified as drivers of tumour growth. An issue with basket trials is that they assume that the heterogeneity between tumour types and location is ancillary to heterogeneity in genetic make-up of the tumour, and that is an untested assumption. Simply put, we test a very particular element of the condition (its genetic make-up) and assume, possibly incorrectly, that this is the key and sole driver of response and progression.
IC5	Often historical cohorts are used to estimate a comparator. What is frequently forgotten is that over time medical care has improved and, hence, that the difference in efficacy is not all attributable to the intervention, while this is often the main assumption of the indirect comparison. Matching strategies are very useful to estimate comparator arms, but often the choice of covariates in the matching model has a sizable impact on the outcome, a source of structural uncertainty that is often not parameterised. One would hope to find a directed acyclic graph to be drawn up to a-priori define relevant covariates for the matching model and solid arguments when other than the pre-defined strategies are used.
IC6	Conducting network-meta analysis will be increasingly complicated in the future due to single arm trials and uniquely defined sub-populations.
IC7	When observational data is used to estimate a control arm, missing data and errors abound. This type of data is often more subjective in nature due to the uncontrolled setting of its collection, which requires intensive collaboration with those who collected the data in the

	first place. Nonetheless, this 'real-world data' provides a better reflection of clinical practice such as the true population in which a drug will be provided and multiple treatment lines. Additionally, observational data, regardless of whether it is a real-world registry or single-arm trial, can provide information that is unlikely obtained through randomised trials; it may provide the synthetic/real-world/historical control arms to single-arm studies of personalised medicine interventions. That is where we see most interest of clinicians.
IC8	An issue with the use of historical cohorts is that data becomes outdated rather fast due to improvements in other treatments than the one studied and improvements in the general health of the population. Also, historical cohorts have not always measured the intermediate endpoints that were central to the trial or have measured them differently.
IC9	Observational data used to assess relative efficacy is worse than RCT data, and perhaps cost-effectiveness models that rely in this type of data should only be approved in research.
IC10	While observational studies better reflect clinical practice than an RCT, it is increasingly difficult to find patients with the specific mutation that is studied in an RCT.
IC11	Historical data can be used to estimate control arms but often a genomic test is new as well, resulting in historical control data to be unadjusted for a specific mutation.
IC9	Limited data often results in using surrogate endpoints and we need evidence on the link between surrogate endpoints and long-term outcomes to be presented when these relationships are assumed in the model
IC13	Perhaps we should not perform HTA altogether when evidence of effectiveness is lacking.
Interventions aiming to cure; methods to compensate for potential bias in the extrapolation of outcomes	
IC1	A cure is clearly defined: eradication of disease with no long-lasting consequences. This definition should not be confused with model results in which individuals are considered to be cured following model assumptions rather than true observations. Often it is assumed that surviving up to a certain point in time in the future can be considered a 'cure', but that is often incorrect and may not align with the underlying interaction between the treatment and the condition. While these cure models may be sensible when we understand the mechanics of the disease and can identify that disease is no longer present and that the risk of relapse is absent, then a cure model may be appropriate, but more often than not this is not known and then assuming a cure seems to be heroic. Perhaps, in interactions with key opinion leaders, we should not ask them if a person is cured, but the probability of living up to a certain point in the future.
IC3	The cure hypothesis should often be testable in historical cohorts in those that responded well to previous treatments. But perhaps the cure hypothesis is more about managed entry agreements than modelling: the assumption can be incorporated in the model, but the managed entry agreement may capture the risk of this assumption being incorrect.

IC5	Often the follow-up in trials is too short to identify someone as being cured. We must rely on the assessments of clinical experts to evaluate the cure hypothesis, or a well described mechanism of the disease that renders the cure hypothesis following treatment likely.
IC6	It is a challenge to demonstrate a 'cure' with trial data. Relying on experts is required in extrapolating treatment effects but other than what is usually done, asking experts to select their preferred extrapolation curve, a more formal process should be used in order to infer the confidence intervals, such as the SHeffield ELicitation Framework (SHELF) method. If experts select their preferred curve, information is lost on the uncertainty in their choices and any parameter uncertainty explored later on can no longer be informed by the uncertainty of the experts. An important piece of information is external evidence: if we have a historical cohort, we can compare predictions of the model with this cohort.
IC7	How a cure is defined depends on the condition and the expected remaining survival. In CAR-T treatment, there is still insufficient data to assess if the condition is cured. Simply choosing a cure model is not a sufficient approach: the model outcomes need to be tested against different parametric models and time horizons.
IC8	Simply put, there is no way to know if you are under- or overestimating long-term outcomes, which is why extensive sensitivity analyses should be conducted including those with an assumption of a waning effectiveness. If waning effectiveness is included, the model should at least consider that clinicians may explore new treatments. Nonetheless, managed entry agreements could be used to mediate the risk of incorrectly assessing the long-term benefit. Regarding cure, it should not be forgotten that patients may experience uncertainty about their disease status at each follow-up.
IC12	There are different scenarios that need to be explored; simply looking at the observed period only, projecting the findings over the lifetime and the definition of a waning period. While expert opinion plays an important role in defining the final model, we have to acknowledge that in situations of uncertainty, 10 guesses might not be much better than one guess. Perhaps the issue of uncertainty is not an issue of health economic modelling. Alternatively, other approaches to getting good value for money could be used such as competitive tendering and pay for performance agreements. The outcomes of these arrangements could even be included in the model as well as the administrative costs of the arrangement itself.
IC11	We should not evaluate PM differently: the issue of extrapolation of results is not unique to this category of treatments. Expert opinion could be elicited to infer opinions on likely scenarios, but it has to be acknowledged that sometimes we simply do not know.
IC9	The term 'cure' implies restoring full health and life expectancy to population levels. The cure assumption should always be tested against different time horizons and the most plausible time horizons need to be informed by both decision-makers preference and expert expectations.

IC13	When long-term effects and safety is unknown, a lifetime time horizon should not be used. Shorter time horizons may be acceptable because a two-year trial may not adequately inform a 30-year forecast.
Deviation from standard discount rates	
IC1	No need to change discount rates, to be consistent across evaluations
IC3	Differential discounting should not be used in an opportunity costs framework
IC6	In PM, costs are often up-front and benefits accrue over time. If managed entry agreements are in place, these should be modelled because costs may fall in different time periods resulting in different incremental cost-effectiveness ratios (ICERs), regardless of differential or uniform discounting
IC7	There is no reason to treat PM treatments differently from others with regard to discounting
IC8	A sensitivity analysis should be conducted to understand the effect of different discounting assumptions
IC12	While there may be arguments for lower discount rates in general, the question is if PM should be evaluated against other discount rates than other therapies. The discount rate used in the economic evaluation should not reward manufactures of PM treatments over others. There are arguments for lower discount rates.(55)
IC11	All interventions should be subjected to the same discounting rules. Otherwise, where do we stop making exceptions? Preferably, the discount rate is based on empirical evidence within the relevant jurisdiction.
IC14	There is no reason to deviate from current discounting practice.
IC13	The same discounting recommendations need to be applied to all interventions.
Uncertainty analysis	
IC9	Uncertainty can be methodological uncertainty, structural uncertainty or parameter uncertainty.
IC1	Uncertainty is not unique to PM. The main difference is that some drivers of uncertainty, such as the duration of treatment effect or the relative effectiveness of treatments are jointly present. Perhaps sometimes it would be easier to present the set of assumptions that have to be made for a therapy to be cost-effective to experts and have them inform you on the likelihood of those assumptions proving to be true simultaneously. A probability sensitivity analysis (PSA) that reflects very large uncertainty may seem to be uninformative and will not tell you what to do, but it at least reflects that uncertainty is very large.
IC3	PM results in more uncertainties as treatment test-combinations inflate the number of strategies as well as multiply these strategies with different types of tests that can be used.
IC4	Eliciting expert opinion is a useful way to dealing with uncertainty but first of all, their input needs to be parametrised for use in the probabilistic uncertainty analysis and second, we

	have to acknowledge that we are asking clinicians for input that they may be rather uninformed about as well.
IC5	Expert elicitation is often used to inform choices that are inherently uncertain. But one expert is not the other. Some standard metrics need to be reported, such as why an individual is an expert and in what, transparent questionnaires used to elicit their opinions or minutes of the meeting. Choices should be traceable to anonymised individuals rather than reporting averages to understand heterogeneity in expert opinion.
IC6	It is always important to include as much uncertainty in the analysis as possible, so value of information methods can be used to prioritise future research, but PM is not fundamentally or conceptually different from other interventions. The evidence may be limited, but the same approaches to identifying uncertainty can be applied. Parameterising structural uncertainty by using different sources all at once can be a relevant approach to dealing with structural uncertainty that may be more informative than scenario analyses (see for example Cope et al. (56)).
IC7	While parameter uncertainty in PM may be similar to other interventions, structural uncertainties abound. But often the impact of important modelling decisions is not evaluated in uncertainty analysis. For example, estimating progression free survival separate from overall survival ignores underlying competing risks. Also often overlooked is behaviour in clinical practice. Even if a biomarker has 100% sensitivity and 100% specificity, patients and clinicians may not actually use that information in their decision making. Hence, the willingness of patients and clinicians to de-escalate treatment based on a biomarker-based risk of progression, for example, should be considered in health economic analysis, since it is not the test result itself that results in different clinical outcomes, but any treatment changes that are made based on test result.
IC8	The timing of effects and costs may warrant sensitivity analysis on discount rates
IC10	The cost of genetic testing is often uncertain and may be rather heterogeneous between labs
IC11	Small sample sizes are a sizeable issue.
IC14	Models that are submitted for reimbursement decisions should adopt standard methods, such as a lifetime time horizon, regardless of uncertainty. But an often-overlooked issue is that PM requires the combination of different data sources, and that each of these data sources has its own types of bias.
IC9	There is uncertainty as to where in the treatment pathway a test is positioned and this should be included, for example by parametrising the structural uncertainty around the test position informed by experts.
IC13	The uncertainty around clinical pathways are especially relevant to PM, but this can be handled with methods as we have now. The model structure should reflect an optimal

	implementation of the test. Modelling the real world (e.g. diminished effectiveness) may confound the ability to detect cost-effectiveness when practice improves.
Values beyond QALYs	
IC1	It is not correct to assume that currently values beyond the QALY do not receive any weight in decision-making. They might not be a formal part of the cost-effectiveness model, but they receive attention and value in deliberations during decision-making.
IC2	Additional elements of value might be captured by the QALY and we do not really know where additional value overlaps with current value
IC3	Some of the elements that are deemed to be values beyond the QALY are probably already captured in the QALY. If there is value that is obviously not captured in, for example, quality of life measurement, a committee will see that and value it. There is also disproportional attention to elements of value as opposed to elements of disutility. For example, knowing you live with an incurable disease may result in a lot of disutility.
IC4	Values beyond the QALY may be useful, but we need to acknowledge that this value is probably also present in interventions foregone, so the opportunity costs need to be adjusted for additional elements of value. We also need to acknowledge that when we incorporate additional elements of value in a cost-effectiveness analysis, that we are effectively willing to trade-off health against something like, for example, the value of hope.
IC5	Cost-effectiveness analysis was always intended to be just a part of health technology assessment. Sometimes HTA and cost-effectiveness analysis overlap almost completely because the cost-effectiveness model captures all relevant elements of value. In other instances, other elements of value need to be weighed in the decision-making process. There is quite a bit of talk about the value of hope, but one has to wonder: should, at equal cost-effectiveness between two treatments, a society be willing to invest societal resources in a treatment that results in slightly more hope than another one? I do not think that generating hope apart from true realised outcomes should be rewarded.
IC6	The 'value flower' that is often discussed has petals that are overlapping -partly due to differences in definitions between researchers- which may cause double counting. With regards to the value of hope: the 'right tail' of the distribution that is considered to be the source of hope is often without any underlying data at the point of the decision, it is based on extrapolations. I am not convinced that should be rewarded by adjusting willingness-to-pay thresholds. Also, the value of hope can be negative for risk-averse individuals. The value of risk-reduction and insurance value can be important in PM. First of all, the whole point is to make sure that the benefits of treatment are stratified to a smaller population in whom greater effectiveness is achieved. At a given value of risk-averseness, this may have a value. Second, it can be worthwhile to model the insurance value for a patient population when there is a treatment available which provides protection against financial and physical

	risk. Finally, it is often forgotten that estimating the value for society is separate from a reimbursement decision which may involve a normative position on extent to which the realised value needs to be reflected in the price.
IC7	The purpose of the QALY metric is to compare across conditions. To allow this, it should not be subjected to a range of elements of value that may differ across analyses. Also, there is more transparency on these elements of value when they are not 'hidden' in the QALY.
IC8	Involving additional elements of value may overcomplicate cost-effectiveness analysis and its communication.
IC10	The value of hope and the reduction of uncertainty can be relevant items that we intend to measure, and we explore ways of adjusting the lambda to reflect additional elements of value.
IC12	Insurance value is not unique to private insurance systems as also collective insurance bodies may not include everything in their basic benefit package. With regards to the value of hope, we need to acknowledge that also many people are disappointed that they are not in the 'right tail' of the distribution, and dashed hope may be worse than not having hope at all. There may be utility in anticipation, for example being protected from human papillomavirus infection in the future, which may be more than simply the avoidance of QALYs lost when the disease would occur.
IC11	First and foremost: costs and QALYs are the main elements of the analysis and additional elements of value may, by and large, may already be captured by the QALY or should not even be considered in the first place. The goal is to maximise population health which is distinct from maximising hope and anticipation. There may be elements of differentiating value against which QALYs can be traded-off, such as severity and equity. The question then becomes if indeed there is a societal preference for trading health or even social welfare against some of the other values that are being discussed.
IC14	Values beyond the QALY should be included in models, but then it should also be taken into account in the valuation of the opportunity costs.
IC13	Incorporating additional elements of value which can be called personal utility or patient preference can be done but should not be incorporated in the ICER to avoid a lack of transparency. Rather, it should be presented separately from the ICER to allow evaluation. Also, additional values can belong to clinicians, patients and decision-makers and should be reported separately.
Values beyond the current patient	
IC1	Values beyond the current patient should be taken into account as interventions may yield effects beyond the patient such as in reducing burden of caregivers. In communicable diseases one also models the effect of treating a patient has on others.

IC5	Costs and QALYs that occur in others than the patient under treatment can be included but should be reported separately for them to be identifiable as separate elements in the evaluation.
IC6	If a patient is cured, parents will have a long period without concern over the health of their child and that is a relevant treatment effect. That does not mean, however, that societies need to be willing to pay extra for achieving that value as there is a difference between assessing value and addressing the reimbursement question. Also, the opportunity costs may also have additional elements of benefit beyond the patient that is currently not included in its estimation.
IC11	If we are going to take into account the benefit of treatments in caregivers and family members, we need to address these values in the estimation of opportunity costs as well.
IC14	These can be taken into account but only if it is expected to change the decision. Otherwise it would be a waste of time and resources.
IC9	Values beyond the patient can be included but should be reported separately
IC13	Values beyond the index patient are important and computationally not difficult to implement but data are often lacking. If cascade testing in family members can avoid adverse outcomes it should be modelled. In some instances, time horizons beyond the current patient's illness can be used.
Perspective	
IC1	The question on perspective is not specifically introduced or answered by new personalised treatments. In fixed health care budgets, the result of the analysis will be the same unless the budget becomes affected by savings in other non-health budgets.
IC5	In some instances, the additional value created beyond health can be so large that they have to be taken into account in the evaluation.
IC7	The issue of perspective is not specific to or related with PM
IC8	For very debilitating diseases which prevent patients from going to school the improved socioeconomic status/productivity gain needs to be taken into account
IC14	Whether or not a societal perspective should be adopted is not specific to PM.
IC13	The societal perspective should always be at least a scenario analysis if possible.
Equity	
IC1	PM, by definition, are a function of one's characteristics and as a consequence one may have a property that societies do not wish to discriminate on that perfectly aligns with probability of treatment effect. Hence, PM can increase equity concerns.
IC3	Authorities that take equity considerations serious should systematically consider equity concerns and if they do so, PM should not pose a threat to increased inequity. The whole point of PM is that individuals respond differently to treatment, so that has to be accepted from the outset.

IC5	Equity issues do not receive equal attention from the different international reimbursement authorities
IC6	The data that underlies economic evaluations may be affected by equity concerns. For example, estimating algorithms to predict treatment efficacy depends on the sufficient representation of equity-relevant groups in the training data
IC7	PM may result in increased inequality, as not everyone will get the same treatment, but its effect on equity considerations differs for each case. Accessibility to genetic tests is relevant and may differ between socio-economic groups.
IC8	Personalised treatments often require quite some knowledge of genetic testing which is often not available in physicians in more geographically remote areas.
IC10	There is a concern that in private insurance settings the results of genetic tests can be used by health insurance companies in manners that was not intended when the test was conducted.
IC9	Higher susceptibility to disease following tests can result in lower admission rates or higher premiums to health insurance.
IC13	Cost-effectiveness analysis can look at subgroups in ethnic populations known to possess higher rates of pathogenic variations to allow the evaluation of equity effects. Inequity can also occur between rural and urban populations when PM technologies are concentrated in urban areas.
Conditionality of test sequences and/or test outcomes	
IC1	The model time should start when the patient undergoes the test. The population that undergoes the treatment is probably smaller than the population that underwent the test, but only looking at the group that tested positive, rather than the suspected group, results in a different scope of analysis. In principle, the costs of the test are borne by the suspected group and will need to be attributed to the treated group. Test itself might also have effects on quality of life that need to be incorporated.
IC7	If multiple tests are deployed, the timing of the test and the number tests that are conducted can be parametrised. Ideally the correlation between tests is taken into account, but this requires patient-level data, which might not always be available.
IC11	Tests can be conducted in different order and that may have treatment consequences. In breast cancer, for example, IHC and FISH can be conducted in parallel or consecutively and this may impact treatment choices. It is not clear cut how this needs to be dealt with, but it seems important to make sure that testing pathways are studies more thoroughly in clinical trials.
IC14	Conditionality of test sequences can be important but also complex to model, so first it has to be identified if it is likely to have a sizeable impact on the ICER. What can be important is to look at the added value of testing larger gene panels. That is often promoted but not

	necessarily a better allocation of resources if single gene panels capture the relevant actionable insights.
IC9	When tests are performed sequentially, and at different time-points, the disease status and the probability of disease changes over time and this needs to be taken into account when possible, but often there is insufficient data.
IC13	Multiple tests are often conducted at once or in series and as a consequence identifying a comparator can be complicated
Time period until start of treatment	
IC6	The model should start when the test procedure starts. An example is CAR-T therapy: often survival functions are applied in the model from receiving treatment onwards, while in reality the time from harvesting T-cells until infusion should also be modelled. Patients may die in the interval from harvesting to infusion and this is an integral part of the analysis.
IC7	The turn-around time of the test should be included. In CAR-T therapy patients have to await the customisation of their treatment and may receive chemotherapy in the meantime, but sometimes they may not survive up to the point in time where they can receive the treatment.
IC12	It is important that the test used in evidence generation and evaluation is the same as the test used in real practice, which may be affected by lab-differences. The costs of the test need to be assessed against their true costs, which often depends on whether a lab conducts the test or the test is commercially marketed. In general, I would not include the costs of downstream tests that follow genomic testing because these tests are conducted for different reasons.
IC11	The time period between testing and receiving treatment may vary and may largely affect results, which is a prominent issue in conditions with high mortality. Using sufficiently small cycle lengths allows taking this issue into account.
IC9	In the evaluation of gene therapies, there has often been a whole diagnostic process in identifying a certain mutation which is often omitted from cost-effectiveness analysis. The question is if the patients are at risk during the waiting period, if that is the case it should be included in cost-effectiveness analyses, and otherwise, it should not be included on the effect side but only on the cost side.
IC13	Including waiting time is very important especially when there is a narrow window for the intervention, such as with pre-school children with neurodevelopmental or other conditions. A diagnosis that comes too late is the same as missed diagnosis.
Patients' willingness to be tested & treatment adherence	
IC6	The hypothesis is that who are tested adhere to treatment better because there is less uncertainty about the probability of success. However, this effect is already accounted for in the assessment of effectiveness. Generally speaking, it is difficult to account for

	adherence separately, as it is unknown to what extent adherence is already accounted for in the assessment of effectiveness.
IC7	For whole genome sequencing it is important to, up front, decide if patients are willing to be exposed to incidental findings wish not to be informed about those findings. We have done research into different services models to implement delivering additional findings in routine practice. From another study by the Melbourne Genomic Health Alliance and from literature, it is known that the moment at which additional findings are offered (e.g. during the first genetic counselling session in which the provision of primary findings is also discussed, or at a later point) impacts the uptake.
IC13	Willingness to be tested, the anxiety of knowing and not knowing, is a part of personal utility.
Clinicians' adherence to protocols and guidelines	
IC7	There is a use for preference studies into the association between the accuracy of a study and the likelihood that a clinician will use the results in treatment decisions in clinical practice.
IC11	Like patients, clinicians may not adhere to the outcome of a test, due to patient or clinician preference.
IC14	The real-world use of tests is often suboptimal. Tests are often only prescribed when common first or second line treatments do not have sufficient effectiveness. Similarly, clinicians do not adhere to the test the way we expect them to and that should be included in the model.
Miscellaneous	
IC3	Meta analyses become increasingly complicated due to PM: treatments can no longer be easily pooled.
IC2	Personalised pathways mean that cohort type models may be less appropriate than microsimulations which may be better in taking into account patient characteristics that drive treatment choices.
IC4	Capacity constraints are a serious concern in PM: tests are not as available as we may think they are.
IC4	We evaluate test-treatment combinations and it is not possible to separate the elements of value that arise from the test from those that follow from treatment.
IC4	Issues in PM can be categorised in methodological issues, such as the appropriate evaluative framework, technical issues, such as programming challenges resulting from complex sequences, practical issues, such as data availability or knowing the unit costs of the test and organisational issues, such as how to pay for test treatment combinations. While there may be practical and organisational concerns, these should not be confused with methodological

	ones or used to push alternative evaluative frameworks for PM compared to other interventions.
IC5	Decision-uncertainty is often on the size of the QALYs gains, not the costs of the tests.
IC5	It is not easy to deal with negative test results, for which the treatment pathway is often diffuse which can mean that a range of comparators needs to be used.
IC7	It is a misunderstanding that discrete event simulation models require a lot of data. They are more flexible models than Markov-type models and is a more natural representation of reality.
IC10	Conducting a genomic test can have large psychological consequences for patients that underwent them and often creates demand for more downstream tests.
IC12	The evaluation of PM is not necessarily different from the assessment of rare diseases or targeted therapies which also suffer from limited data and small patient population. The main issue of PM is that all the limitations occur simultaneously.
IC12	When the European Medicines Agency (EMA) approves treatments based on limited evidence as is currently the case, we can count on having less RCT data in the future to base our evaluations on.
IC12	Perhaps commissioning trials is a better way forward than working with observational data gathered from conditional reimbursement.(57)
IC11	The ability to test someone does not mean that all geographic areas have the capacity, budget or knowledge to act on the result of the test.
IC14	The variation of treatments and treatment lines in oncology should be better captured in models
IC9	It is up to the decision-maker to decide if the cost of the test should be included in the cost-effectiveness analysis.
IC9	The question on the distribution of reward between the test and the treatment is different from the question a cost-effectiveness analysis seeks to address.
IC13	The low prevalence of a mutation in a population is a specific challenge of PM as it hinders getting good epidemiological data.
IC13	In the case of hereditary conditions and cascade testing, clinicians have a duty of care that transcends the index patient.

Summarised responses for the evaluators of health economic models

INTERVIEWEE	PARAPHRASED RESPONSE
Effectiveness data from non-randomised (controlled) studies	
IC15	It is no longer realistic to assume that the reimbursement decision can follow directly after the market approval decision, due to the tendency in both the US Food and Drug Administration (FDA) and EMA to allow conditional market approval on lower quality evidence, evidence that cannot inform reimbursement decisions. The issue is not personalised medicine itself, but the fact that evidence with more uncertainty (e.g. due to small RCTs or single arm studies) is often accepted for certain health technologies.
IC16	Regulatory bodies only assess the risk-benefit relationship, while in HTA you have to take into account the existing treatments and cost-effectiveness. Regulatory and HTA agencies have different mandates.
IC17	The solution to poor quality data is not in more sophisticated modelling methods, there simply should be higher demands on manufacturers. Given the difference in demands from regulators and reimbursement agencies, manufacturers should be more aware that market approval is only first step and that further studies/data might be needed for reimbursement.
IC18	Decision-makers simply have to accept observational data to inform decisions when no other data is around.
Interventions aiming to cure; methods to compensate for potential bias in the extrapolation of outcomes	
IC15	The time horizon needs to be relevant to the decision problem, hence when interventions have a long-term effect, the time horizon should be lifetime.
IC17	There is no reason to truncate the time horizon, as there is sufficient reason to assume that effects may be long lasting. This may be uncertain, but not a reason to change the time horizon.
IC18	When the cure assumption is uncertain, rules of thumb, such as a maximum time horizon of the median follow-up times two can be used. The definition of a cure is often not very clear.
Deviation from standard discount rates	
IC15	Discounting effects different from costs means implicitly assuming the that cost-effectiveness threshold would go up in the future.
IC17	Differential discount rates can be good to address the uncertainty in effects, for example through using a higher discount rate on health effects.
Uncertainty analysis	
IC15	Uncertainty is something that decision makers can deal with. A situation where there is low uncertainty about the added value of a medicine is still to be preferred, as decision

	makers would then not be uncertain about what they should be paying. However, high decision uncertainty can be managed to a certain extent through negotiating a lower price/discount.
IC16	It is not possible to solve uncertainty about long-term effectiveness associated with gene therapies by collecting more data, because it simply has not been observed yet. A product that has less uncertainty should be valued more than the same product with more uncertainty. Uncertainty is not be taken into account by decision makers enough currently. Uncertainty should be reflected in the price or willingness to pay threshold.
IC17	Uncertainty can be dealt with in probabilistic uncertainty analysis. When the uncertainty has been identified, it can be dealt with through for example lower prices or using different discount rates for health effects (which will also mean that a lower price is needed).
IC18	We should embrace uncertainty and be open about it. Personalised medicine does not require more uncertainty analyses than are already conducted, but we should be very transparent about the amount of uncertainty underlying a model.
Managed entry agreements	
IC15	The single-point-in-time reimbursement decision following market approval has had its time and may need to be replaced by a dynamic decision system, possibly using managed entry agreements.
IC16	Paying per QALY afterwards would take care of uncertainty, but it hands over the whole surplus to the manufacturer completely. Other way of treating uncertainty is collecting evidence over time and adjusting payment over time. But it is still difficult to prevent that the whole consumer surplus is handed over to the manufacturer.
IC17	Observational registries are expensive and those who have experience with running them are not necessarily positive about it. One may favour managed entry agreements in general, but there may be legal restrictions in a country limiting its enforcement. A higher risk premium is often a better alternative.
IC18	To make better use of the potential of managed entry agreements, those who deal with them should have a good understanding of the health economic model and the drivers of costs and uncertainties in it.
Values beyond QALYs	
IC15	There may be values beyond the QALY, but this does not mean that a reimbursement authority should be willing to pay for those values; optimising values beyond the QALY might not be an authority's objective.

IC16	It is plausible that the QALY does not capture all relevant values. Regardless of which value, it is more transparent to deal with additional values outside the QALY, such as through threshold weighing. Scientific spill over is not a value that ought to be included.
IC17	Some of the values beyond the QALY are already taken into account in the deliberation procedure following the health economic assessment. Several European jurisdictions already take 'severity' into account. The value of hope or scientific spill-over should not be included. Regardless of which values, it is preferable to not include these values in the QALY but adjust the willingness to pay threshold where relevant.
IC18	A decision-maker would be happy to consider additional values beyond the QALY if there is a sufficient evidence base. This could include values beyond the current patient.
Perspective	
IC17	A general issue with the societal perspective is that it, to some extent, discriminates against people who can't or don't work. There are many tricky ethical questions associated with these types of evaluation.
IC18	The societal perspective can be adopted if the societal gains included have sufficient evidence base.
Equity	
IC16	The problem is not that some subgroups might not get a treatment based on it not being effective, but there is an equity issue if a health economic model identifies that a treatment is not cost-effective in a certain sub-group and that it therefore would not be reimbursed in that subgroup.
IC17	Equity issues are not specific to personalised medicine
IC18	When personalised medicines are specialised and cost-intensive, there is no equal access between regions which may increase inequalities.
Modelling tests	
IC15	If the testing is routine practice its costs do not need to be included in the economic evaluation as it is not part of the decision-problem.
IC16	If the test is related to the treatment, then the costs of the test should be included. But the test can also be part of standard care for other reasons, then it should not be included.
IC17	The utility impact of not being among those who can benefit from new treatment is not something that should be included
Imperfect implementation	
IC15	Models should reflect reality: if drug-holidays are a thing, the model should reflect that.
IC16	Models can be used to identify subgroups where the treatment is most (cost-) effective.
IC17	The initial requirement should be that a model reflects perfect implementation, but it could be validated if that turns out to be true in the long run.

IC18	A decision-maker needs to be informed about both the optimal and the real world scenario.