|  |  |
| --- | --- |
| **Test description (Please document how the test is conducted, as well)** | **Expected result of the test** |
| *Pre-analysis calculations* | |
| Does the technology (drug/device, etc.) acquisition costs increase with higher prices? | Yes |
| Does the drug acquisition cost increase for higher weight or body surface area? | Yes |
| Does the probability of an event, derived from an odds ratio (OR)/ relative risk (RR) / hazard ratio (HR) and baseline probability, increases with higher OR/RR/HR? | Yes |
| If survival parametric distributions are used in the extrapolations, can the formulae used for the Weibull (generalized gamma) distribution generate the values obtained from the exponential (the Weibull or Gamma) distribution(s) under some parameter transformations? | Yes |
| In a partitioned survival model, does the progression free survival curve or the time on treatment curve crosses the overall survival curve? | No |
| If survival parametric distributions are used in the extrapolations or time-to-event calculations, can the formulae used for the Weibull (generalized gamma) distribution generate the values obtained from the exponential (the Weibull or Gamma) distribution(s) after replacing/transforming some of the parameters? | Yes |
| Is hazard ratio calculated from Cox proportional hazards model applied on top of the parametric distribution extrapolation found from the survival regression? | No, it is better if the treatment effect that is applied to the extrapolation comes from the same survival regression in which the extrapolation parameters are estimated. |
| For the treatment effect inputs, if the model uses outputs from WINBUGs, are the OR, HR and RR values all within plausible ranges? (should be all non-negative and the average of these WINBUGs outputs should give the mean treatment effect) | Yes |
| *Event-state calculations* | |
| Calculate the sum of the number of patients at each health state | Should add up to the cohort size |
| Check if all probabilities and number of patients in a state are greater than or equal to zero | Yes |
| Check if all probabilities are smaller than or equal to one | Yes |
| Compare the number of dead (or any absorbing state) patients in a period with the number of dead (or any absorbing state) patients in the previous periods? | Should be larger |
| In case of lifetime horizon, check if all patients are dead at the end of the time horizon | Yes |
| *Discrete event simulation specific:* sample one of the “time to event” types used in the simulation from the specified distribution. Plot the samples and compare the mean and the variance from the sample | Sample mean and variance & the simulation outputs should reflect the distribution it is sampled from. |
| Set all utilities to one  Set all utilities to zero | The QALYs accumulated at a given time would be the same as the life years accumulated at that time  No utilities will be accumulated in the model |
| Decrease all state utilities simultaneously (but keep event based utility decrements constant) | Lower utilities will be accumulated each time |
| Set all costs to zero | No costs will be accumulated in the model at any time |
| Put mortality rates to 0 | Patients never die |
| Put mortality rate extremely high | Patients die in the first few cycles |
| Set the effectiveness, utility and safety related model inputs for all treatment options equal | Same life years and QALYs should be accumulated for all treatment at any time |
| In addition to the inputs above, set cost related model inputs for all treatment options equal | Same costs, life years and QALYs should be accumulated for all treatment at any time |
| Change around the effectiveness, utility and safety related model inputs between two treatment options | Accumulated life years and QALYs in the model at any time should be also reversed |
| Check if the number of alive patients estimate at any cycle is in line with general population life table statistics | At any given age, the % alive should be lower or equal in comparison to the general population estimate |
| Check if the QALY estimate at any cycle is in line with general population utility estimates | At any given age, the utility assigned in the model should be lower or equal in comparison to the general population estimate |
| Set the inflation rate of the previous year higher | The costs (which are based on a reference from previous years) assigned at each time will be higher |
| Calculate the sum of all ingoing and outgoing transition probabilities | Both should be one |
| Calculate the number of patients entering and leaving a tunnel state throughout the time horizon | Numbers entering = Numbers leaving |
| Check if the time conversions for probabilities were conducted correctly. | Yes |
| *Decision tree specific:* calculate the sum of the expected probabilities of the terminal nodes | Should sum up to one |
| *Patient-level model specific:* check if common random numbers are maintained for sampling for the treatment arms? | Yes |
| *Patient-level model specific:* check if correlation in patient characteristics is taken into account when determining starting population? | Yes |
| Increase the treatment acquisition cost | Costs accumulated at a given time will increase during the period when the treatment is administered |
| *Population model specific:* set the mortality and incidence rates to zero | Prevalence should be constant in time |
| *Result calculations* | |
| Check the incremental life years and QALYs gained results. Are they in line with the comparative clinical effectiveness evidence of the treatments involved? | If a treatment is more effective, it generally results in positive incremental LYs and QALYs in comparison with the less effective treatments |
| Check the incremental cost results. Are they in line with the treatment costs? | If a treatment is more expensive, and if it does not have much effect on other costs, it generally results in positive incremental costs. |
| Total life years > total quality adjusted life years | Yes |
| Undiscounted results > discounted results | Yes |
| Divide undiscounted total QALYs by undiscounted life years. | This value should be within the outer ranges (maximum and minimum) of the all utility value inputs. |
| Subgroup analysis results: How do the outcomes change if the characteristics of the baseline change? | Better outcomes for better baseline health conditions and worse outcomes for worse health conditions are expected. |
| Could you generate all the results in the report from the model (including the uncertainty analysis results)? | Yes |
| Does the total life years, QALYs and costs decrease if a shorter time horizon is selected? | Yes |
| Is the reporting and contextualization of the incremental results correct? | The use of the terms such as: “dominant”/ “dominated”/ “extendedly dominated”/ “cost-effective” etc. should be in line with the results.  In the incremental analysis table involving multiple treatments, ICERs should be calculated against the next non-dominated treatment. |
| Are the reported ICERs in the fully incremental analysis non-decreasing? | Yes |
| If disentangled results are presented, do they sum up to the total results? (e.g. different cost types sum up to the total costs estimate) | Yes |
| Check if half cycle correction is implemented correctly (total life years with half cycle correction should be lower than without) | The half cycle correction implementation should be error free. Also check if it should be applied for all costs, for instance if a treatment is administered at the start of a cycle, half cycle correction might be unnecessary. |
| Check the discounted value of costs/QALYs after 2 years | Discounted value=undiscounted/(1+r)2 |
| Set discount rates to zero | The discounted and undiscounted results should be the same |
| Set mortality rate to zero | The undiscounted total life years per patient should be equal to the length of the time horizon |
| Put the consequence of adverse event/discontinuation to zero. (zero costs and zero mortality/utility decrements) | The results would be the same as the results when AE rate is set to zero. |
| Divide total undiscounted treatment acquisition costs by the average duration on treatment. | This should be similar to treatment related unit acquisition costs |
| Set discount rates to a higher value | Total discounted results should decrease |
| Set discount rates of costs/effects to an extremely high value | Total discounted results should be more or less the same as the discounted results accrued in the first cycles |
| Put adverse event/discontinuation rates to zero and then to extremely high level. | Less costs higher QALYS/LYs when adverse event rates are 0, higher costs and lower QALYS/LYs when AE rates are extreme |
| Double the difference in efficacy and safety between new intervention and comparator and report the incremental results. | Approximately twice of the incremental effect results of the base case. If this is not the case : report and explain the underlying reason/ mechanism |
| Do the same for a scenario in which the difference in efficacy and safety is halved. | Approximately halve of the incremental effect results of the base case. If this is not the case : report and explain the underlying reason/ mechanism |
| *Uncertainty analysis calculations* | |
| Are all parameters subject to uncertainty included in the one-way sensitivity analysis (OWSA)?  Check if the OWSA includes any parameters associated with joint uncertainty (e.g. parts of a utility regression equation, survival curves with multiple parameters). | Yes  No |
| Are the upper and lower bounds used in the one-way sensitivity analysis used confidence intervals based on the statistical distribution assumed for that parameter?  Are the resulting ICER, incremental costs/QALYs with upper and lower bound of a parameter plausible and in line with a priori expectations? | Yes  Yes |
| Check that all parameters used in the sensitivity analysis have an appropriate associated distributions  - upper and lower bounds should surround the deterministic value (i.e. Upper bound ≥ mean ≥ Lower bound)  - standard error and not standard deviation used in sampling  - Lognormal / gamma distribution for hazard ratios and costs/ resource use  - Beta for utilities and proportions/probabilities  - Dirichlet for multinomial  - Multivariate normal for correlated inputs (e.g. survival curve or regression parameters)  - Normal for other variables as long as samples don’t violate requirement to remain positive when appropriate | Yes |
| Check PSA output mean costs, QALYs and ICER compared to the deterministic results. Is there a large discrepancy? | No (in general) |
| If you take new PSA runs from the excel model do you get similar results? | Yes |
| Is(are) the CEAC line(s) in line with the CE scatter plots and the efficient frontier? | Yes |
| Does the PSA cloud demonstrate an unexpected behavior or has an unusual shape? | No |
| Is the sum of all CEAC lines equal to 1 for all WTP values? | Yes |
| Are the explored scenario analyses provide a balanced view on the structural uncertainty? (i.e. not always looking at more optimistic scenarios) | Yes |
| Are the scenario analysis results plausible and in line with a priori expectations? | Yes |
| Check the correlation between 2 PSA results (i.e. costs/QALYs under the SoC and costs/QALYs under the comparator) | Should be very low (very high) if different (same) random streams are used for different arms |
| If a certain seed is used for random number generation (or previously generated random numbers are used), check if they are they scattered evenly between 0-1 when they are plotted? | Yes |
| Compare the mean of the parameter samples generated by the model against the point estimate for that parameter, use graphical methods to examine distributions, functions | The sample means and the point estimates will overlap, the graphs will be similar to the corresponding distribution functions (e.g. Normal, Gamma, etc.) |
| Check if sensitivity analyses include any parameters associated with methodological/ structural uncertainty (e.g. annual discount rates, time horizon). | No |
| Value of information analysis if applicable: Was this implemented correctly?  Which types of analysis? Were aggregated parameters used? Which parameters are grouped together? Does it match the write-up’s suggestions?  Is EVPI larger than all individual EVPPI?  Is EVPPI for a (group of) parameters larger than the EVSI of that (group) of parameter(s)?  Are the results from EVPPI in line with OWSA or other parameter importance analysis (e.g. ANCOVA)? | Yes |
| Did the electronic model pass the black-box tests of the previous verification stages in all PSA iterations and in all scenario analysis settings? (additional macro can be embedded to PSA code, which stops the PSA when an error such as negative transition probability, is detected) | Yes |
| Check the correlation between 2 PSA results (i.e. costs/QALYs under the SoC and costs/QALYs under the comparator) | Should be very low (very high) if different (same) random streams are used for different arms |
| OWSA=one-way sensitivity analysis; ICER = incremental cost-effectiveness ratio; PSA = probabilistic sensitivity analysis; WTP = willingness to pay; CE = cost-effectiveness; CEAC = cost-effectiveness acceptability curve; LY = life years; QALYs = Quality adjusted life years; OR = odds ratio; RR= relative risk; HR = hazard ratio | |