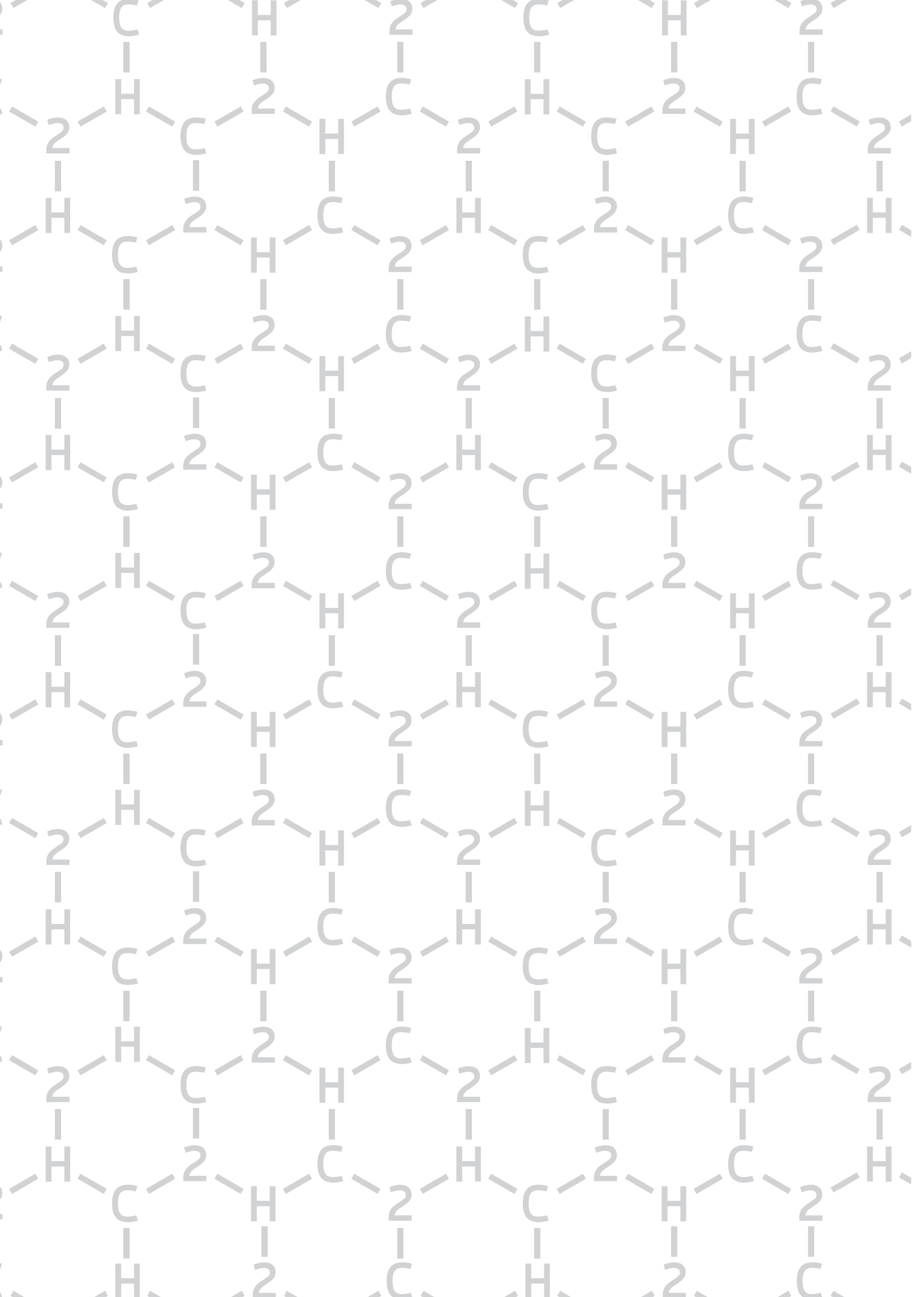




Guideline for Preparing Cost-Effectiveness Evaluation to the Central Social Insurance Medical Council

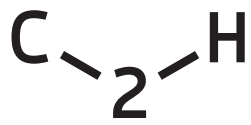
—
Version 3.0 approved by CSIMC on 19th January, 2022

The English version is a translation of the original version in Japanese. The Japanese version is preferentially applied in cases of discrepancy between the two versions.



Contents

1	Objectives	4
2	Analysis perspective	5
3	Target population	6
4	Comparator	7
5	Additional benefits	8
6	Methods of analysis	11
7	Time horizon	12
8	Choice of outcome measure	13
9	Sources of clinical data (excluding costs)	15
10	Calculation of healthcare costs	16
11	Public long-term care costs and productivity loss	19
12	Discounting	21
13	Modeling	22
14	Uncertainty	23
	Terminology	24
	Abbreviations	29

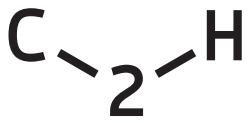


1 Objectives

- 1.1 This guideline presents standard methods to perform cost-effectiveness evaluations of medicines and medical devices selected by the Central Social Insurance Medical Council (“selected technologies”).
- 1.2 This guideline is applied to manufacturers’ submissions and academic analysis (review and re-analysis).

2 Analysis perspective

- 2.1 The perspective of the analysis should be specified. In particular, the analysis should consider the range of costs corresponding to this perspective.
- 2.2 “Public healthcare payer’s perspective” is a standard. It uses costs, comparator, and target populations to reflect the situation of public healthcare insurance in Japan.
 - 2.2.1 Even when an analysis is conducted from a perspective other than the “public healthcare payer’s perspective,” an analysis from the “public healthcare payer’s perspective” should be submitted.
 - 2.2.2 There are some healthcare technologies that are not covered by public healthcare insurance but are publicly funded, such as some prophylactic procedures (e.g., health check-ups, vaccinations). Analyses including these technologies should be submitted from the “public healthcare payer’s perspective.”
- 2.3 If the effect on public long-term care costs is important with regard to the selected technology, it is acceptable to perform an analysis from the “public healthcare and long-term care payer’s perspective.”
- 2.4 If the introduction of a selected technology influences productivity, it is acceptable to perform an analysis that includes productivity loss from the broader perspective.



3 Target population

- 3.1** Patient populations that meet the indications when the target technology is selected should be considered as the target population of the cost-effectiveness evaluation.
 - 3.1.1** In the case that a new indication is approved between the selection of target technology and the determination of the framework of analysis, it is to be included in the target population.
 - 3.1.2** A new indication is added after the time defined by item 3.1.1 and a new evaluation is performed after the first evaluation is completed if it may influence the results.
- 3.2** An analysis should be conducted for each population if the technology has multiple indications or subpopulations that differ in outcome, application method/dose, and administration or comparator.
 - 3.2.1** However, if item 3.2 is difficult to achieve, it is acceptable to perform analyses of limited population(s) considering factors such as the number of patients or features of the illness. The exemption is determined based on agreement between the manufacturer and the National Institute of Public Health/public academic group in consultation.
- 3.3** The percentage of each patient group should be determined in a manner that reflects the recent clinical context of the target population.

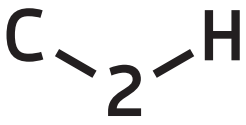
4 Comparator

- 4.1** The comparator should be principally selected from technologies that are widely used in clinical practice and are expected to be replaced by the selected technology when it is introduced to treat the target population. Among them, technologies that result in better outcomes should be selected.
- 4.1.1** Non-treatment or watchful waiting can also be used as comparators.
- 4.1.2** Except for the cases described in item “4.1.1,” the comparator should be selected from technologies that can be used by public healthcare insurance.
- 4.1.3** If a single comparator cannot be determined based on item “4.1,” the comparator should be selected by considering the comparators in randomized controlled trials (RCTs), similar technologies for the official pricing, and cost-effectiveness based on agreement after consultation with C2H.
- 4.2** Sufficient explanation of the reasons underlying the selection of the comparator is needed.

5 Additional benefits

- 5.1** When a cost-effectiveness evaluation is conducted, whether the additional benefit of the selected technology to the comparator is proven should first be evaluated.
- 5.2** Evaluations of the additional benefit should be conducted on the basis of a systematic review (SR) of RCTs compared with a technology selected in section “4.” If appropriate, the results of unpublished clinical studies/trials may also be included in the SR.
- 5.2.1** When an SR is conducted, research questions (RQs) should be clearly presented. For example, a definition of structured RQs according to PICO (P: patient, I: intervention, C: comparator, O: outcome) may be provided.
- 5.2.2** There may be technologies with a similar action mechanism or function category to the selected technologies or comparator(s) determined in section “4” that are expected to show equivalent outcomes. These technologies can be included as an intervention (I) or comparator (C) in the SR if deemed appropriate after consultation with C2H.
- 5.2.3** For outcome (O) in item “5.2.1,” the most appropriate clinical outcomes (e.g., a “true outcome”) should be selected from among clinical effectiveness, safety, and health-related quality of life (HRQOL).
- 5.2.4** Descriptions of the inclusion/exclusion criteria, databases used, search algorithm, and research selection process (inclusion flow diagram) are required in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement.
- 5.2.5** It is acceptable to utilize any existing reliable SR. In such cases, the existing review will be used solely or in combination with a new additional study. In this case, it should confirm the consistency of the existing review by considering the RQs and coverage of the most recent literature.

- 5.2.6** If appropriate, pooled results via a meta-analysis should be presented. In such cases, the required reporting factors include the statistical method, assessment of heterogeneity, forest plot, pooled results, and confidence interval.
- 5.2.7** When it is obvious that no RCTs have been performed, the process described in section “5.2” can be skipped upon agreement in consultation.
- 5.2.8** A time point between determining the framework of analysis and manufacturer’s submission can be used as a cut-off date for the literature search in the SR.
- 5.2.9** There may be cases in which the results of new clinical trials are published after the cut-off date defined in item “5.2.8” but are regarded as important information for the cost-effectiveness evaluation (e.g., clinical trials with a large sample size or reliable results that differ from current studies). Inclusion of these trials in the SR should be considered. In that case, additional SR is not required.
- 5.3** When no studies are available based on the result of SR described in section “5.2,” the additional benefit is evaluated by SR of comparative non-RCT (e.g., observational) studies based on section “5.2.” In that case, sufficient explanation regarding the research quality is needed (e.g., study design, differences in background factors between groups, statistical analysis, sample size, and number of institutions).
- 5.4** When more reliable results are obtained, the additional benefit can be evaluated via re-analysis of existing observational study or registry data, if agreed upon in consultation. In that case, sufficient explanation on the research quality is needed (e.g., study design, differences in background factors between groups, methods of statistical analysis, sample size, and number of institutions).
- 5.5** When there are no RCT studies using the same comparator selected in section “4” but there are RCT studies of the selected technology compared to others, the additional benefit is evaluated via an indirect comparison using the SR results.



- 5.6** When only single-arm clinical studies of selected technologies are available, an indirect comparison should be performed based on the SR results of the selected technologies and the comparator.
- 5.7** When an indirect comparison is performed, the following items should be considered.
- 5.7.1** When individual patient data are available, the difference of background factors should be adjusted using an appropriate method such as MAIC (matched adjusted indirect comparison).
 - 5.7.2** When individual patient data are not available, an adjusted indirect comparison using RCT or network meta-analysis should be used.
 - 5.7.3** When neither individual patient data nor results of RCTs are available, a naïve indirect comparison may be acceptable if other methods cannot be used. In such a case, the uncertainty of the results should be carefully considered.
 - 5.7.4** If an indirect comparison is conducted, sufficient explanation on the prerequisites for the indirect comparison (e.g., heterogeneity of illness, severity, and patient background or similarity of the studies) is also needed.
- 5.8** There may be cases in which the results obtained by the methods in items “5.3” to “5.7” have serious problems regarding the quality of the studies. However, it is expected that the selected technology is not inferior to the comparator. In such cases, the analysis described in section “6” can be performed, assuming the outcome of the selected technology is equivalent to that of the comparator.
- 5.9** When there are no clinical data available on the selected technology in humans, the analysis described in section “6” can be performed, assuming the outcome of the selected technology is equivalent to that of the comparator(s), if appropriate. This is based on considering the approval of the Pharmaceuticals and Medical Devices Agency (PMDA).

5.10 When the SR results obtained by the methods in items "5.2" to "5.7" show that the outcomes of the selected technology are inferior to those of the comparator(s), no cost-effectiveness analysis is performed.

6 Methods of analysis

- 6.1** A cost-effectiveness analysis should be used. In this guideline, cost-effectiveness analysis is defined as an analysis that calculates cost and effectiveness separately without converting effectiveness into monetary units.
- 6.2** If the analysis described in section “5” reveals an additional benefit, the incremental cost-effectiveness ratio (ICER) should be calculated based on the expected cost and effectiveness in each treatment group.
- 6.3** In the following cases, only the expected cost and effectiveness in each group need to be presented and the ICER should not be calculated.
- 6.3.1** In cases where the technology has better effectiveness and lower costs relative to the comparator, the technology is considered “dominant” without a calculation of ICER.
- 6.3.2** A cost comparison with the comparator (a so-called “cost minimization analysis” [CMA]) should be performed if the analysis described in section “5” does not demonstrate an additional benefit but the outcome of the selected technology appears to be equivalent to that of the comparator. In such a case, the result is either “cost saving” or “cost increase.”
- 6.4** If the selected technology has multiple indications or sub-populations defined in item “3.2” and/or “3.2.1,” ICER should be calculated for each indication or sub-population.
- 6.5** If a cost-effectiveness analysis for a selected technology published in an academic journal or an evaluation from a health technology assessment (HTA) agency is available, these results should also be presented.

7 Time horizon

- 7.1 The time horizon should be sufficiently long to evaluate the influence of the technology on cost and effectiveness.
- 7.2 The same time horizon should be applied for both cost and effectiveness.
- 7.3 The reason for setting this time horizon should be specified.

8 Choice of outcome measure

- 8.1 Quality-adjusted life years (QALY) should be used in principle.
 - 8.1.1 When it is difficult to calculate QALY, and CMA is applied, other outcome measures can be used, if appropriate.
- 8.2 When QALY is calculated, the QOL score should be measured by a preference-based measure (PBM) in principal.
 - 8.2.1 If Japanese QOL scores (utilities) are newly collected for a cost-effectiveness analysis, EQ-5D-5L is recommended as the first choice.
 - 8.2.2 If data corresponding to item "8.2" are unavailable, it is acceptable to use mapping of other appropriate HRQOL data. When using a QOL score obtained from mapping, the conversion into a QOL score via an appropriate method should be explained.
- 8.3 When the QOL score is assessed by PBM, the subjects' own QOL responses should be used.
 - 8.3.1 In the case of using PBMs, responses from a proxy (e.g., family member or caregiver) may be used only when the subject cannot respond.
 - 8.3.2 In the case of using PBMs, proxy responses from a healthcare professional should be avoided due to possible discrepancies with subjects' own responses.
 - 8.3.3 If it is difficult to directly collect QOL scores from patients, it is acceptable for general people to evaluate the presented health scenario by standard gamble (SG), time trade-off (TTO), and discrete choice experiment (DCE). However, QOL scores measured via these methods are largely influenced by the presented health scenario. Therefore, the limitations of this method should be considered; for example, the same QOL score should be used for the same health state.

- 8.4** As long as a QOL score that satisfies items "8.2" and "8.3" is available, the use of Japanese results is preferentially recommended.
- 8.4.1** If Japanese research is absent or insufficient but high-quality research is available overseas, it is acceptable to use the data collected overseas.
- 8.5** In the case of the analysis from the public healthcare and long-term care payer's perspective, the QOL scores' influence on the informal caregiver may be considered if actual data exist.

9 Sources of clinical data (excluding costs)

- 9.1** ICER calculations should preferentially use effectiveness, safety, and QOL data (inclusion of parameters such as transition probability for model analysis) derived from high-quality research with a high evidence level reflective of practical clinical results in Japan.
- 9.1.1** The selection of effectiveness, safety, and QOL data on the basis of an SR of all the clinical research is recommended. This review may also include unpublished clinical study/trial data if appropriate.
- 9.1.2** Data with a high evidence level should be used preferentially. The use of data deemed appropriate from the viewpoints of research quality, target population, and external validity is recommended (for example, it is possible that the results of an RCT may differ markedly from practical clinical results).
- 9.1.3** Data derived from re-analysis of existing study and/or registry data can be used if appropriate. In this case, detailed information on patient background, statistical methods, etc. must be provided.
- 9.2** Japanese data should be used preferentially if there is evident heterogeneity between Japanese and overseas data.
- 9.3** If the data do not differ statistically significantly between the selected technology and the comparator(s), pooled data of both groups should be applied. Otherwise, when considering factors such as data or rationale that support the difference, effect size (clinical meaning), it should be explained that the difference is interpretable.
- 9.4** Regarding evaluation of medical devices, if there are reliable and quantitative data, analysis reflecting “learning effect” (i.e., improvement of treatment effect via the accumulation of clinicians’ experience) or “product improvement effect” can be submitted in addition to analyses not considering the effects, upon agreement in consultation with C2H.

10 Calculation of healthcare costs

- 10.1 Only public healthcare costs should be included in the case of analysis from public healthcare payers' perspective.
- 10.2 The healthcare costs of each health state include only related costs that are directly affected by the selected technology, and do not include unrelated costs.
- 10.3 The healthcare costs of each health state should reflect the average resource consumption and standard clinical practices in Japan.
- 10.4 It is recommended that claims databases established in Japan, which reflect actual clinical practice from the viewpoint of item "10.3," should be used to estimate the costs of each health state if appropriate. However, this recommendation does not apply to cases in which it is difficult to define health states using only information from claims data or if there are insufficient data in the database.
 - 10.4.1 A definition of each health state and its rationale is required when claims data are used for cost estimation.
 - 10.4.2 The methods and rationale for the method to estimate costs (including handling outliers and unrelated costs) should be shown.
- 10.5 Micro-costing (by medical fee schedule etc.) based on the definitions of the standard clinical process can be used if it is difficult to estimate the costs of each health state via a claims database or if micro-costing is more appropriate.
 - 10.5.1 In the case of the application of micro-costing, the rationale for costing should be shown from the viewpoint of item "10.2." It may be better to identify relevant items and/or estimate the amount of medical resource consumption in the claims database.

- 10.5.2 When micro-costing is used, the medical resource consumption and unit costs should be reported separately.
- 10.5.3 In principle, for the estimation of resource consumption in item “10.5.2,” the resource consumption of injection products should be calculated by the number of vials rather than by patient dosage.
- 10.6 The estimation should include not only the costs of the selected technology and the comparator(s) but also the costs of adverse events and related future events etc.
- 10.7 An analysis of public healthcare costs should include not only the portion of costs paid by the insurer but also those paid by the government and patients as co-payment (i.e., the total public healthcare expense).
 - 10.7.1 Based on the principal in item “2.2.2,” the analysis should include the costs of health check-ups, vaccinations, or similar procedures that are funded publicly but not reimbursed by Japan’s public healthcare insurance.
- 10.8 Unit costs should be derived from the latest medical fee schedule, drug price lists, or similar resources. It is particularly essential to use the latest unit costs for the selected technology or comparator(s).
 - 10.8.1 Even if existing cost-of-illness studies or analyses of claims data are used, unit costs at the time of evaluation rather than when the medical resources are consumed should be applied. It is acceptable to make adjustments such as multiplication by the medical payment system revision rate.
 - 10.8.2 Such adjustments may be omitted if the influence on results is minimal.

- 10.9** If generics of the comparator(s) are already on the market, an analysis using these costs should] also be submitted.also submitted.
- 10.10** Even if the costs of selected technology and/or comparator(s) are included in a bundled payment, the estimation should be based on fee-for-service payment.
- 10.11** Future costs should also be estimated based on current medical resource consumption and unit costs.
- 10.12** Calculations of medical resource consumption based on other countries' data will require attention to possible differences in healthcare technology use between Japan and other countries. The unit costs in Japan should be applied in the analysis.

11 Public long-term care costs and productivity loss

- 11.1** An analysis from the perspective of the public healthcare and long-term care payer may include public long-term care costs in addition to base-case analysis. It is acceptable to include public long-term care costs in additional analyses only if they can be estimated by Japanese data.
- 11.2** When public long-term care costs are included in the analysis, it is recommended that these costs be calculated based on the care level.
- 11.3** The amount utilized under public long-term care insurance should be based on the actual resource consumption. If this consumption is difficult to determine, it is acceptable to use the average amount utilized per beneficiary or similar data.
- 11.4** An analysis including productivity loss can be additionally performed on top of the base-case analysis. However, judgments regarding the appropriateness of including productivity losses should consider the possibility of working in the context of the illness characteristics. It is acceptable to include productivity losses in additional analyses only if they can be estimated by Japanese data.
- 11.5** Decreases in productivity losses may be classified as follows:
- (A) those arising directly from healthcare technology (e.g., treatment-related shortening of hospital stay) and
 - (B) those arising indirectly from outcome improvements (e.g., alleviation of illness, survival period extension).

When productivity loss is included in an analysis, only (A) should be included in the cost calculation.

- 11.6** Productivity losses should be estimated using the human capital method. This method estimates the loss via the expected earned wage in the absence of illness.

- 11.6.1** The unit wage used for estimations of productivity loss should be the average wage across all industries, all ages, and both genders or the average wage for each age group in all industries and both genders derived from the latest “Basic Survey on Wage Structure” (Wage Census) and not discriminate by income.
- 11.6.2** Estimations of productivity loss require an actual investigation into the employment status in the target population (i.e., a measure of the days or hours of work missed). The actual measured number of days or hours should then be multiplied by the average wage across all industries, all ages, and both genders to accurately estimate the productivity loss.
- 11.6.3** If the item described in item “11.6.2” is difficult to perform, productivity loss should be calculated by multiplying the expected number of days (excluding holidays) or hours of work missed multiplied by the average wage across all industries, all ages, and both genders. A 100% employment rate should be assumed for those aged 18 years and older. However, note that this method may overestimate productivity losses.
- 11.7** If other individuals (e.g., family members) experience productivity losses due to the provision of nursing or informal patient care, it is acceptable to count these productivity losses as costs under the same conditions and using the same methods as those used to calculate the patient’s productivity loss.
- 11.7** Time costs that are unrelated to a decrease in work should not be included in the cost estimations.

12 Discounting

12.1 Future costs and effectiveness must be discounted and converted into present values.

12.1.1 Discounting is not needed if the time horizon is one year or less or is otherwise sufficiently short to ignore the influence of discounting.

12.2 Both cost and effectiveness should be discounted at a rate of 2% per year.

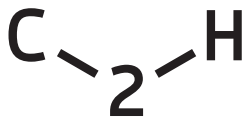
12.3 The discount rate should be subjected to sensitivity analysis and be changed at the same rate of 0-4% per year for both cost and effectiveness.

13 Modeling

- 13.1** To predict prognosis and future expenses, it is acceptable to conduct a model analysis using a decision analytic model, the Markov model, and/or other models in accordance with the principle described in section "7."
- 13.2** The model analysis should present the validity of the model, including the following.
- (A) Internal validity: this addresses why a model with a given structure has been created, whether the natural history of illness has been sufficiently evaluated, and whether the parameters used are appropriate.
 - (B) External validity: this addresses whether the estimation yielded from the model is appropriate in comparison to other existing clinical data.
- 13.3** The assumption used to create the model should be clearly described.
- 13.4** All parameters and data sources used for model creation should be presented.
- 13.5** The model should be submitted in the form of electronic files. The model must be easily understood by third-party experts and all main parameters (transition probability, QOL score, and healthcare costs) must be alterable.
- 13.5.1** It is ideal that not only total costs but also the breakdown (in the case of micro-costing, the medical resource consumption and unit costs of each item) can be changed. Especially, the unit costs of the selected technology and comparator(s) must be adjusted for each academic analysis group.
- 13.6** Half-cycle correction should be used in the Markov model if the length of the Markov cycle is long and its influence on the results is not negligible.

14 Uncertainty

- 14.1 If the analysis setting has multiple scenarios and this could affect the results, a scenario analysis should be conducted.
- 14.2 For situations in which the uncertainty is high because of a long time horizon, a shorter-term analysis is necessary, such as an analysis of the period for which clinical study data are available.
- 14.3 If no available studies involve a comparison with the comparator according to section “5,” particularly when a comparison has been made concerning results between single-arm studies, a sensitivity analysis with a sufficiently wide range is required because of the large degree of uncertainty.
- 14.4 Sensitivity analyses are needed for parameters with large variances, those based on assumptions rather than actual data, those with possible heterogeneity between domestic and other countries’ data, and others.
- 14.5 When the variance of the estimator should be considered (parametric uncertainty), the range moving parameter in the sensitivity analysis can refer to the 95% confidence interval of the estimator.
- 14.6 A probabilistic sensitivity analysis (PSA) is also desirable. In such cases, the distribution used for analysis, scatter plots of the cost-effectiveness plane, and cost-effectiveness acceptability curves (CEAC) must be presented.



Terminology

Additional benefit

In a cost-effectiveness analysis, the additional benefit relative to the comparator should be demonstrated before calculating the ICER. The endpoint of effectiveness used to demonstrate the additional benefit does not always need to be equal to the outcome used for the cost-effectiveness analysis but should be clinically significant. If additional benefit is judged to be shown, cost-effectiveness analysis should be performed. On the other hand, if no additional benefit is shown, cost of both treatment should be compared by so called "CMA".

Cost-effectiveness analysis

Economic evaluations of healthcare technologies are often divided into the following patterns: (a) cost-minimization analysis (CMA), in which the outcome is deemed equivalent and only cost is analyzed; (b) cost-effectiveness analysis (CEA), which uses outcome units other than QALY (LY, event avoidance, etc.); (c) cost-utility analysis (CUA), which uses QALY; and (d) cost-benefit analysis (CBA), which involves an evaluation of outcomes after conversion into monetary units.

However, CMA, CEA, and CUA can all be considered analogous in situations where the cost and outcome are estimated in different units. For this reason, these types of analysis are collectively termed "cost-effectiveness analyses" in this guideline.

Discounting

In a cost-effectiveness analysis, a discount at a constant rate is usually made to convert future costs and arising (or obtained) outcomes to current values. Costs converted to the current value after applying yearly discounts (C_p) can be calculated from the cost at i years later (C_i) and the discount rate (d) using the following equation:

$$C_p = \frac{C_i}{(1+d)^{i-1}}$$

The same calculation can be used for effectiveness.

Dominant/dominated

If a technology is lower in cost and equivalent or higher in effectiveness than the comparator is, the technology is called “dominant.” If the technology is higher in cost but equivalent or lower in effectiveness relative to the comparator, the technology is called “dominated.”

Evidence level

Diverse classification methods for evidence levels are available. MINDS (Medical Information Network Distribution Service) set forth the following classification:

- I Systematic review/meta-analysis of RCTs
- II From one or more RCTs
- III From a non-randomized controlled study
- IVa Analytical epidemiological study (cohort study)
- IVb Analytical epidemiological study (case-control or cross-sectional studies)
- V Descriptive study (case reports or series)
- VI Views of an expert committee or individual experts that are not based on patient data

However, it has been often noted that the results of experimental studies such as randomized controlled trials (RCTs) can differ from real-world clinical data. Economic evaluations of healthcare technologies should primarily use data with a high level of evidence, although consideration should be given to appropriate clinical data.

Human capital method

The “human capital method” is used to estimate productivity loss based on the wages originally expected to be earned. However, when viewed from a long-term standpoints, the inability of an individual to work does not always lead to a productivity loss because in a situation with an employment rate less than 100%, as other individuals are likely to work in place of the affected

individual who is unable to work. For this reason, one view suggests that productivity losses should include only friction costs (e.g., based on the period needed to restore the initial production level). Wages should be originally estimated through an investigation of the period for which an individual was actually unable to work because of illness. If this estimation is difficult due to lack of data including housework, it is acceptable to set the employment rate at 100%. From the viewpoint of fairness, the mean wage across all industries, all ages, and both genders should be used as the unit wage, regardless of the actual unit wage for individuals.

Incremental cost-effectiveness ratio

The incremental cost-effectiveness ratio (ICER) is the incremental cost divided by the incremental effectiveness. ICER of treatment A compared with B is calculated using the following equation:

$$ICER = \frac{IC}{IE} = \frac{C_A - C_B}{E_A - E_B}$$

IC: incremental cost

E_A: expected effectiveness

IE: incremental effectiveness

of treatment A

C_A: expected cost of treatment A

E_B: expected effectiveness

C_B: expected cost of treatment B

of treatment B

ICER is an indicator of the cost to acquire one unit of effectiveness. A lower ICER indicates higher cost-effectiveness.

Indirect comparison

When clinical studies yield results for “A vs. B” and “A vs. C,” an estimation of the results for “B vs. C” in which no direct comparison is available from the head-to-head results is called an “indirect comparison.” If no head-to-head study involving an appropriate comparator is available, an indirect comparison may occasionally be used.

The following conditions must be satisfied to enable indirect comparison: the results for “A vs. B” must also be applicable to the population “A vs. C” and the results for “A vs. C” must also be applicable to the population “A vs. B.” This is called an “assumption

of similarity.” When an indirect comparison is performed, it is necessary to test this assumption and to use appropriate statistical methods (for example, adjusted indirect comparison rather than naïve indirect comparison). This approach also enables analyses based on more advanced methods such as network meta-analyses (or multiple treatment comparisons; MTCs).

Mapping

When preference-based measure-determined QOL scores are unavailable, it is sometimes advantageous to use PRO data to calculate the QOL score used for cost-effectiveness analysis. The conversion of scores between measures is called “mapping.” Mapping is acceptable as a second-best method when no other data are available but should be performed only after sufficient assessment of the statistical validity.

Meta-analysis

Meta-analysis is a method by which the results from a systematic review are integrated statistically to yield integrated values or their confidence intervals. If the heterogeneity is small, a fixed-effect model is usually used. If the heterogeneity is large, random-effect or Bayesian models are usually employed. The results are often depicted as forest plots. If a comparison is made among multiple treatments rather than between two treatments (pairwise comparison), a “network meta-analysis” is used, employing different methods (ref. “Indirect comparison”).

Probabilistic sensitivity analysis

Probabilistic sensitivity analysis (PSA) is a technique used to determine the distributions of incremental cost, incremental effectiveness, and ICER by applying model parameters to the distribution. The results of a PSA are usually shown as a scatter plot on the cost-effectiveness plane and as a cost-effectiveness acceptability

curve (CEAC), defined as $f(\gamma) = \Pr(\gamma \cdot IE - IC > 0)$ (IC: incremental cost, IE: incremental effectiveness, γ : willingness to pay).

Productivity loss

Depending on the perspective, a loss resulting from the inability to perform work/housework because of illness (or benefit from early recovery) may be counted as a cost (i.e., productivity loss) but is not included in the base-case analysis. It is acceptable to consider not only the loss experienced directly by the patient but also losses experienced by family members or others arising from the need to provide nursing or informal care. According to this guideline, however, an indirect productivity loss resulting from an improvement in the patient’s health states (e.g., survival period extension) is not included in productivity loss to avoid double counting (i.e., counting a factor as both effectiveness and costs). Only a productivity loss directly attributable to the healthcare technology (e.g., shortened hospital stay) is permitted for inclusion.

Quality-adjusted life year

A quality-adjusted life year (QALY) value is calculated by multiplying the life years (LYs) by the QOL score. A QOL score of 1 indicates full health, whereas 0 indicates death. If an individual has survived for two years under a health states with a QOL = 0.6, the LY is two years and the QALY is $0.6 \times 2 = 1.2$ (equivalent to 1.2 years survival under full health). If the QOL score changes over time, the QALY is represented by the area under the curve of the QOL score over time, as illustrated in the figure below.



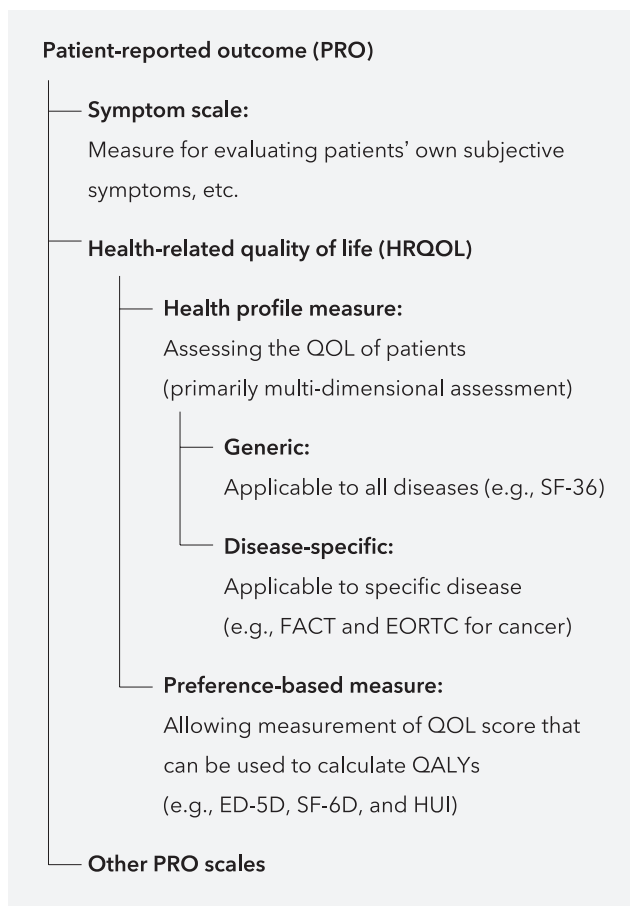
Quality of life (QOL) score (utility)

The health states (i.e., value obtained from the health states) is scored using a one-dimensional scale ranging from 0 (death) to 1 (full health). Negative scores, reflective of a health states “worse than death,” are also possible.

QOL scoring methods can be categorized as follows: (1) direct methods that evaluate health states under a hypothetical situation (or about himself/herself), including the standard gamble (SG) and time trade-off (TTO) methods, and (2) indirect methods that calculate QOL scores from patients’ responses to QOL questionnaires using a scoring algorithm.

The QOL score used for cost-effectiveness analysis cannot always be calculated from any patient-reported outcome (PRO) or QOL data. Cost-effectiveness analysis can utilize only QOL scores determined using a preference-based measure developed for QALY calculation, as described below.

The EQ-5D (EuroQoL 5 dimension) is one currently available measure for which a scoring algorithm has been developed in Japan.



Sensitivity analysis

When uncertainty is present, its influence on the results can be evaluated by changing the parameter in a “sensitivity analysis.” Sensitivity analyses can be further classified as a one- (only one parameter is changed) and two-dimensional (two parameters are simultaneously changed) sensitivity analyses, as well as PSA (simultaneous uncertainty in multiple parameters; see “Probabilistic sensitivity analysis”).

Systematic review

A systematic review is a method by which the literature is comprehensively searched regarding a specific topic and the results are evaluated/reported without bias if at all possible. This method was defined by MINDS as follows: “When defined from the aspects of practical actions, systematic review means ‘searching studies on a given clinical question comprehensively, grouping studies of identical quality on each research design and analyzing/integrating them being accompanied by evaluation of biases.’”

A systematic review is often confused with meta-analysis. The results of a systematic review do not always require statistical integration; this type of systematic review is also known as a “qualitative systematic review.” In cases where the integration of results is deemed appropriate, a meta-analysis of the systematic review results is needed.

Regarding the reporting style for a systematic review (meta-analysis), the style presented in the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement has been used as a standard and can be used as a reference.

Uncertainty

Various types of uncertainty accompany cost-effectiveness analyses.

Broadly, heterogeneity is a type of uncertainty that indicates a situation lacking uniformity in terms of the comparator, healthcare

patterns, targeted patients, and other factors. This differs from the uncertainty in the narrow sense, as explained below. This is not a technical problem related to statistics or health economics but rather arises from real-world variety. If such heterogeneity is present, a sensitivity analysis based on multiple scenarios is recommended.

Uncertainty in the narrow sense can be divided into (a) model and (b) parameter uncertainties. Model uncertainty can result from (a)-1 methodological uncertainty and (a)-2 model structure/assumptions.

Methodological uncertainty, mentioned in (a)-1, arises from the theoretical impossibility of setting uniform methods for the estimation of the discount rate and productivity loss, measuring the QOL score, and other parameters. To avoid this type of uncertainty, it is important to conduct an analysis in accordance with common and standard procedures. If results such as the discount rate are markedly affected, uncertainty should be evaluated through one-way sensitivity analysis.

Uncertainty arising from the model structure/assumption, as mentioned in (a)-2, is caused by the method used to model the health states and treatment processes, selection of parameters for incorporation into the model, assumptions regarding predictions of long-term prognosis beyond the observation period, and other factors. This uncertainty should be evaluated in a sensitivity analysis.

Parameter uncertainty, as mentioned in (b), arises from uncertainty inherent in the parameter estimation. For example, if 10 of 100 subjects develop events during a clinical study, the true incidence rate might not be $10/100 = 0.1$ in the whole population. To deal with this type of uncertainty, which is attributable to statistical inference, it is useful to conduct a PSA in addition to a deterministic sensitivity analysis.

These unrelated costs are not included in the cost.

Unrelated medical costs

Medical costs can be divided into related (i.e., those directly affected by the selected technology) and unrelated (i.e., those affected indirectly through survival extension or those not related to the illness). For example, a hypertension treatment that reduces the incidence of cardiovascular disease and stroke will extend life expectancy, possibly leading to an increase in unrelated medical costs (e.g., costs related to dementia, diabetes, and hemodialysis).

Abbreviations

CBA: Cost-benefit analysis	MTC: Multiple treatment comparison
CEA: Cost-effectiveness analysis	PBM: Preference-based measure
CEAC: Cost-effectiveness acceptability curve	PRO: Patient-reported outcome
CMA: Cost-minimization analysis	PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
CSIMC: Central Social Insurance Medical Council (Chuikyo)	PSA: Probabilistic sensitivity analysis
CUA: Cost-utility analysis	QALY: Quality-adjusted life year
DCE: Discrete choice experiment	RCT: Randomized controlled trial
EQ-5D: EuroQol 5 dimension	RQ: Research question
HRQOL: Health-related quality of life	SG: Standard gamble
ICER: Incremental cost-effectiveness ratio	SR: Systematic review
MAIC: Matched adjusted indirect comparison	TTO: Time trade-off

