

CORE2 HEALTH

---

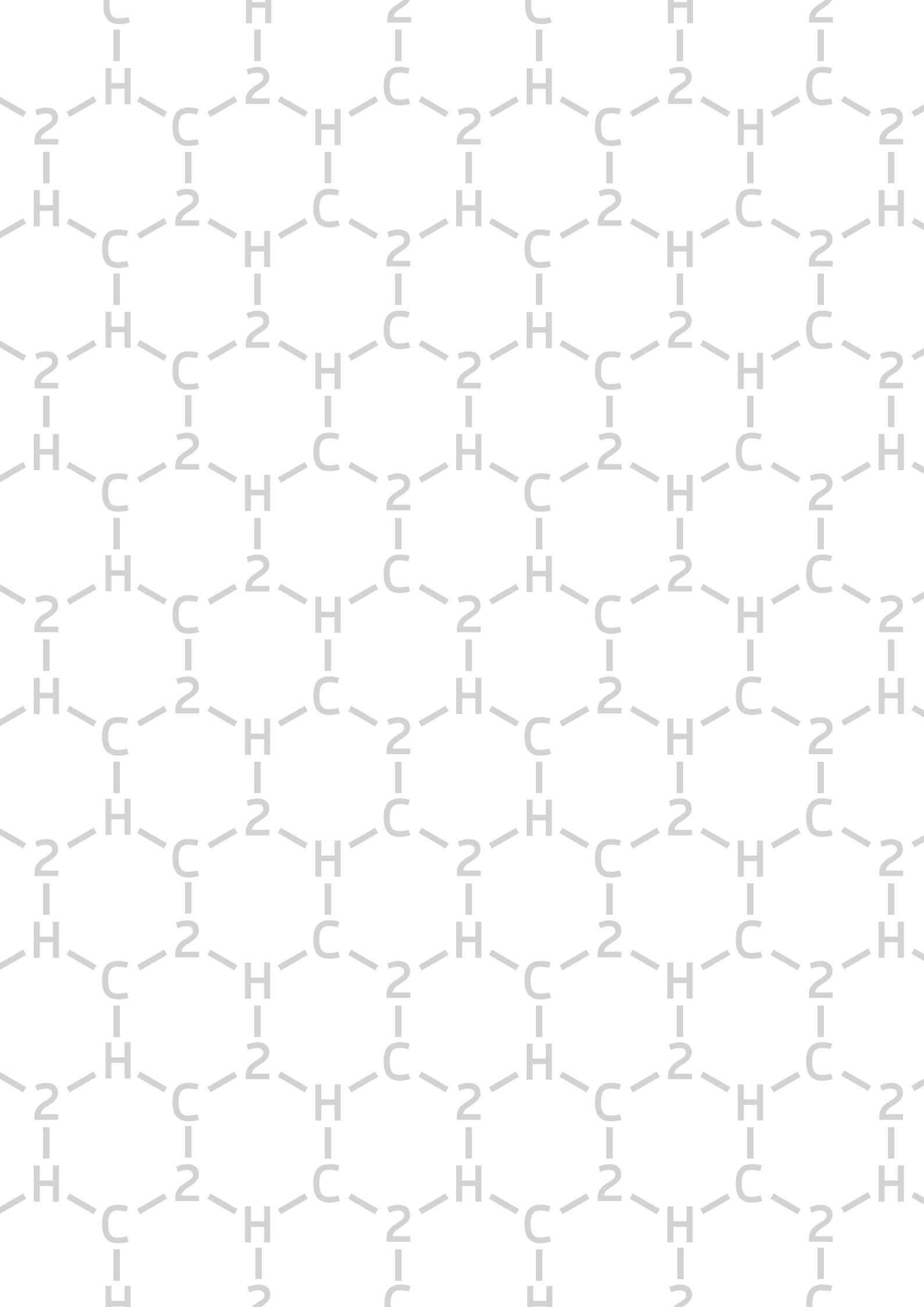
# Guideline for Preparing Cost-Effectiveness Evaluation to the Central Social Insurance Medical Council

—  
Version 2.0 approved by CSIMC on 20th February, 2019

---

Prepared by: Research team (Team Leader: Takashi Fukuda) on cost-effectiveness evaluation supported by Health and Labour Science Research Grants (Strategic Integrated Scientific Research Project)

—  
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.

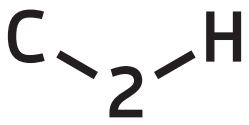


---

**Contents**

---

1	Objectives	4
2	Analysis perspective	5
3	Target population	6
4	Comparator(s)	7
5	Additional benefit	8
6	Methods of analysis	11
7	Time horizon	12
8	Choice of outcome measure	13
9	Sources of clinical data (except 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 perspective that pertains to factors such as costs, comparator(s), and target populations within the range of the 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 also be submitted.
  - 2.2.2 There are some healthcare technologies that are not covered by the public healthcare insurance but are publicly funded, such as some prophylactic procedures (e.g., health checkups, 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 has a direct influence on productivity, it is acceptable to perform an analysis that considers the broader costs and counts productivity loss as a cost.

### 3 Target population

**3.1** Patient populations that meet the indications for the selected technology when the manufacturer's analysis is performed should be considered the target population of the cost-effectiveness evaluation.

**3.1.1** In the case that a new indication (or addition of a new dose and administration) is approved between the time of the selection of target technology and the manufacturer's submission of analysis, the new indication (or dose and administration) is also included in the target population.

**3.2** If the technology has multiple indications or even in single-indication subpopulations which differ in outcome, application method/dose and administration or comparator of cost-effectiveness evaluation, an analysis should be conducted for each population in principal.

**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.

---

## 4 Comparator(s)

- 4.1** The comparator(s) should be principally selected from among technologies which are expected to be replaced by the selected technology at the time when the technology was introduced to treat the target population. Among them, technologies which are widely used in clinical practice and which result in a better outcome 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”, comparator(s) should be selected from among technologies reimbursed by public healthcare insurance.
- 4.1.3** If single comparator cannot be determined based on item “4.1”, the comparator(s) should be selected considering the comparators in randomized controlled trials (RCTs), referred technology when determining the official price, cost-effectiveness, and other factors, based on agreement in consultation.
- 4.2** Sufficient explanation of the reasons underlying the selection of the comparator(s) is needed.

---

## 5 Additional benefit

- 5.1** When a cost-effectiveness evaluation is conducted, the additional benefit of the selected technology to the comparator(s) should first be evaluated.
- 5.2** Evaluations of the additional benefit should be conducted on the basis of a systematic review (SR) of RCTs. The RCTs should be directly compared with the technology selected in section “4.”. The results of unpublished clinical studies/trials may also be included in the SR if they are deemed appropriate.
- 5.2.1** When an SR is conducted, clinical questions (CQs) should be clearly presented. For example, a definition of structured CQs according to PICO (P: patient, I: intervention, C: comparator, O: outcome) may be provided.
- 5.2.2** There may be technologies with similar action mechanism or function category to the selected technologies or comparator(s) determined in section “4.”, which will be expected to show equivalent outcomes to them. These technologies can be included as an intervention (I) or comparator (C) in the SR if they are deemed appropriate in consultation.
- 5.2.3** As outcome (O) in item “5.2.1”, the most appropriate clinical outcomes (e.g. a “true outcome”) should be used to evaluate selected technology from the viewpoint of clinical effectiveness, safety, and health-related quality of life (HRQOL).
- 5.2.4** A description of the inclusion/exclusion criteria, databases used, search algorithm, and research selection process (inclusion flow of information) is 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 directly or in combination with a new additional study. In this case, it



should confirm the consistency of the existing review by considering the CQs and coverage of the most recent literature.

- 5.2.6** If deemed appropriate, pooled results by meta-analysis should be presented. In such cases, the required reporting factors include the employed statistical method, assessment of heterogeneity, forest plot, pooled results, and confidence interval, among others.
  - 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 cost-effectiveness evaluation (e.g., clinical trials with large sample size or reliable results different from current knowledge). Inclusion of these trials in the SR should be considered. In that case, additional SR is not necessarily required.
- 5.3** When no studies or only insufficient studies are available based on the result of SR described in section “5.2”, additional benefit is evaluated by SR of comparative non-RCT (e.g. observational) studies based on section “5.2”, if agreed upon in consultation. In that case, sufficient explanation on research quality is needed (e.g., study design, differences in background factors between groups, methods of statistical analysis, sample size, and number of institutions).
- 5.4** When more reliable results are obtained, additional benefit can be evaluated by re-analysis of existing observational study and/or registry data, if agreed upon in consultation. In that case, sufficient explanation on 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, additional benefit is evaluated by indirect comparison using SR results, if agreed upon in consultation.
- 5.5.1** The applicability of item “5.5” depends on the quality of study on the indirect comparison. 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.
- 5.6** When there are only single-arm clinical studies of selected technologies, SR results of the selected technologies (and comparator(s) if needed) should be shown.
- 5.6.1** In such cases, the evaluation of additional benefit has to consider a number of factors such as the characteristics of the technology and/or disease, background of the participants, and the quality of the studies. Therefore, whether an additional benefit is shown is judged by agreement in consultation.
- 5.7** There may be cases in which the results obtained by the methods in sections “5.3” to “5.6” have serious problems regarding the quality of the studies and 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 selected technology is equivalent to that of the comparator(s).
- 5.8** When there are not any available clinical data 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) and considering the approval of the Pharmaceuticals and Medical Devices Agency (PMDA), if agreed upon in consultation.
- 5.9** When results obtained by the methods in sections “5.2” to “5.6” show that outcomes of the selected technology are inferior to that 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.” allows a judgment that reveals additional benefit, the incremental cost-effectiveness ratio (ICER) should be calculated from 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 is equivalent or superior in terms of effectiveness (non-negative incremental effectiveness) and lower in terms of cost, relative to the comparator, the technology is considered “dominant” without a calculation of ICER.
  - 6.3.2** A cost comparison with the comparator (so-called “cost minimization analysis” [CMA]) should be performed if the analysis described in section “5.” does not demonstrate an additional benefit compared to the comparator, but the outcome of the selected technology appears equivalent to that of the comparator.
- 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 major public health technology assessment (HTA) agency are available, these results should also be presented.

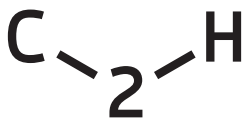
## 7 Time horizon

- 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, upon agreement in consultation.
  
- 8.2** When QALY is calculated, the QOL score should be reflective of the value in a general population (using preference-based measure [PBM] or direct methods such as the standard gamble [SG] and the time trade-off [TTO]). However, systematic difference may exist between QOL scores measured by SG and by TTO.
  - 8.2.1** If Japanese QOL scores are newly collected for a cost-effectiveness analysis, the use of PBMs with a value set developed in Japan using TTO (or mapped onto a TTO score) 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 by proxy responses from a healthcare professional, possible discrepancies from subjects’ own responses should be explained.
  - 8.3.3** If it is difficult to directly collect QOL scores from subjects, it is acceptable for general people to evaluate the presented health scenario by direct methods. It is better for the validity of the presented scenario to be confirmed by clinicians. In this case, use of the TTO method is recommended.



**8.4** As long as a QOL score that satisfies sections “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.

---

## 9 Sources of clinical data (except costs)

- 9.1** Calculations of the ICER 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 Japanese and overseas clinical research is recommended. This review may also include unpublished clinical study/trial data if deemed 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 by re-analysis of existing study and/or registry data can be used if deemed appropriate. In that 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, the rationale that additional benefit is shown by the process described in section “5.” etc. must be explained.
- 9.4** Regarding evaluation of medical devices, if there are reliable and quantitative data, analysis reflecting “learning effect” (i.e., improvement of treatment effect by the accumulation of clinicians’ experience) or “product improvement effect” can be submitted in addition to analysis not considering the effects, upon agreement in consultation.

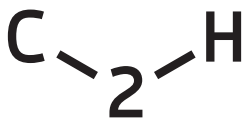
---

## 10 Calculation of healthcare costs

- 10.1** Only public healthcare costs should be included in the case of analysis from the public healthcare payers' perspective.
- 10.2** 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** 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 section "10.3", should be used to estimate the costs of each health state, if deemed appropriate. However, this recommendation does not apply to cases in which it is difficult to define health states using only information from claims data, insufficient data have been accumulated in the database, and so on.
- 10.4.1** 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 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 by 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 "10.3". 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 amount of injection products consumed should be defined by the number of vials rather than by patient dosages.
- 10.6** The estimation should include not only the costs of the selected technology and the comparator(s) but also the costs of factors such as adverse events and related future events.
- 10.7** An analysis of the 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 copayment (i.e., the total public healthcare expenses).
- 10.7.1** Based on the principal in item “2.2.2”, the analysis should include the costs of health checkups, vaccinations, or similar procedures that are funded publicly and 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, not at the time that the medical resources 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, analysis using these costs should be also submitted.



- 10.10** If the costs of selected technology and/or comparator(s) are included in bundled payment, the estimation should be based on fee-for-service payment.
- 10.11** Future costs should also be estimated on the basis of current medical resource consumption and unit costs.
- 10.12** Calculations of medical resource consumption based on overseas data will require attention regarding possible differences in healthcare technology use between Japan and overseas countries. The unit costs in Japan should be applied in the analysis.

---

## 11 Public long-term care costs and productivity loss

**11.1** Public long-term care costs and productivity losses arising from an inability to perform work should not be included in the base-case analysis.

**11.1.1** It is acceptable to include public long-term care costs and productivity losses in additional analyses only if they can be estimated by Japanese data. However, judgments regarding the appropriateness of including productivity losses should consider the possibility of working in the context of the illness characteristics.

**11.2** When public long-term care costs are included in the analysis, it is recommended that these costs should be calculated based on the care level.

**11.3** The amount utilized under public long-term care insurance should be based on the actual quantity of resources consumed. If this quantity is difficult to determine, it is acceptable to use the average amount utilized per beneficiary or similar data.

**11.4** Decreases in productivity losses may be classified as follows:

(A) Decreases arising directly from healthcare technology (e.g., treatment-related shortening of hospital stay);

(B) Decreases 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 calculation of costs.

**11.5** Productivity losses should be estimated using the human capital method. This method was designed to generate estimations based on the expected earned wage in the absence of illness.

**11.5.1** 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.

- 11.5.2** Estimations of productivity loss require an actual investigation of 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 estimate the productivity loss.
- 11.5.3** If the item described in item “11.5.2” is difficult to perform, productivity loss should be calculated by multiplying the expected number of days (excluding holidays) or hours of work missed in the target population 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.
- 11.6** 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.

---

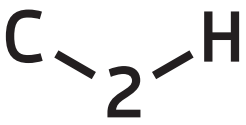
## 12 Discounting

**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 should be changed at the same rate of 0-4% per year for both costs and effectiveness.



---

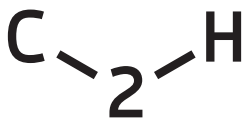
## 13 Modeling

- 13.1** To predict prognosis and future expenses, it is acceptable to conduct a model analysis using a decision analytic model, Markov model, and/or other models in accordance with the principle described in section “7.”.
- 13.2** Model analysis should present the validity of the model. For example:
- (A) Internal validity: This addresses why a model with a given structure has been created, whether or not the natural course of illness has been sufficiently evaluated, whether or not the parameters used are appropriate, and other factors.
  - (B) External validity: This addresses whether the estimation yielded from the model is appropriate in comparison to existing other clinical data, and other factors.
- 13.3** The assumption used to create the model should be described clearly.
- 13.4** All parameters and data sources used for model creation should be shown.
- 13.5** The model used and the calculation processes 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 able to be changed.
- 13.5.1** It is better 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 able to be changed by academic analysis group in the model.
- 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 patterns of clinical practice or other factors are not uniform and this discrepancy could affect the results, analyses based on multiple scenarios should be conducted.
- 14.2 For situations in which the uncertainty is large 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 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 overseas and domestic 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 standpoint, 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  
IE: incremental effectiveness  
C<sub>A</sub>: expected cost of treatment A  
C<sub>B</sub>: expected cost of treatment B  
E<sub>A</sub>: expected effectiveness of treatment A  
E<sub>B</sub>: expected effectiveness of treatment B

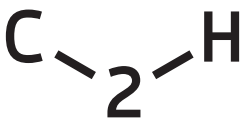
ICER is an indicator of the cost to acquire one unit of effectiveness. A lower ICER indicates higher cost-effectiveness.

---

## 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,  $\gamma$ : 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.



---

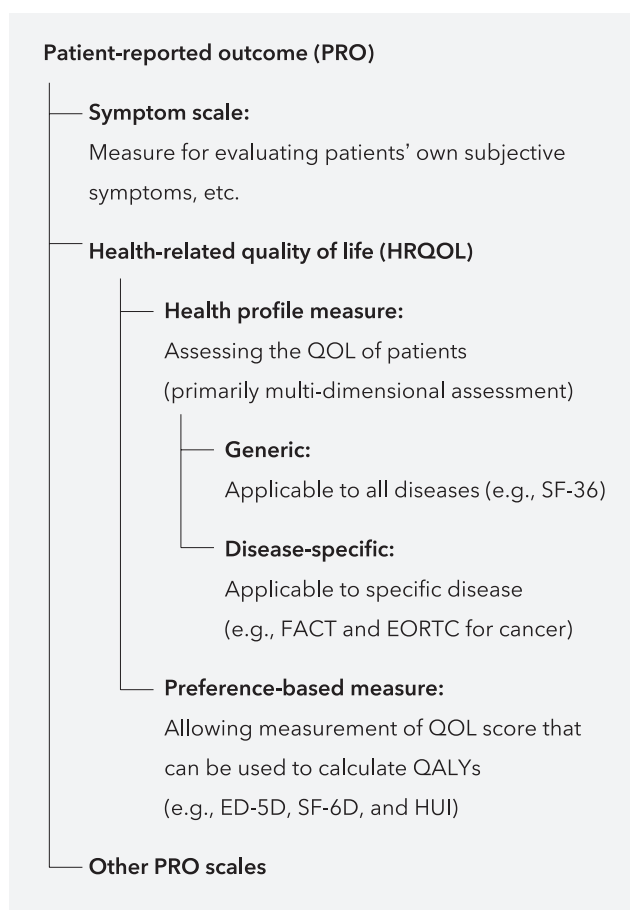
## Quality of life (QOL) score

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.

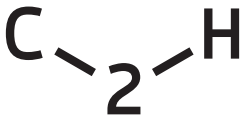
Regarding the reporting style for a systematic review (meta-analysis), the style presented in the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement has been used as a standard and can be used as a reference.

---

## 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.

---

## 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). These unrelated costs are not included in the cost.

---

## Abbreviations

**CBA:**

Cost-benefit analysis

**CEA:**

Cost-effectiveness analysis

**CEAC:**

Cost-effectiveness acceptability curve

**CMA:**

Cost-minimization analysis

**CSIMC:**

Central Social Insurance Medical Council  
(Chuikyo)

**CUA:**

Cost-utility analysis

**CQ:**

Clinical question

**EQ-5D:**

EuroQol 5 dimension

**HRQOL:**

Health-related quality of life

**ICER:**

Incremental cost-effectiveness ratio

**MTC:**

Multiple treatment comparison

**PBM:**

Preference-based measure

**PRO:**

Patient-reported outcome

**PRISMA:**

Preferred Reporting Items for  
Systematic Reviews and Meta-Analyses

**PSA:**

Probabilistic sensitivity analysis

**QALY:**

Quality-adjusted life year

**RCT:**

Randomized controlled trial

**SG:**

Standard gamble

**SR:**

Systematic review

**TTO:**

Time trade-off



