## Annex 3. Fact sheet examples (A. Cost–effectiveness and B. Budget impact)

* + 1. **Cost–effectiveness information sheet for rapid global HIV, viral hepatitis and STI priority-setting**

**Why it matters**

Cost–effectiveness analysis explores which interventions deliver the greatest health benefit to populations within a budget or resource constraint. It is commonly used for determining how to use public funds for health system investment and spending. This is especially important in contexts with limited budgets, in which difficult choices may have to be made.

Cost–effectiveness is not the same as choosing the intervention with the lowest cost. Rather, it identifies which options offer the highest value in terms of providing the best health for the resources available.

**What is cost–effectiveness?**

* Cost–effectiveness compares the additional cost of an intervention to the additional health benefit it delivers.
* The results are summarized using an incremental cost–effectiveness ratio (ICER).
* The ICER is usually expressed as the cost per DALY averted, representing a reduction in population disease burden.
* The ICER is a comparison to alternatives such as the current standard of care or no intervention.

The following considerations apply to ICER estimates:

* defining the comparator and intervention and considering how services are, or would be, delivered for the population in need;
* choosing a perspective (such as health system or societal) that determines what costs and benefits are included;
* using summary health outcomes such as DALYs or quality-adjusted life-years, which combine mortality and morbidity, and for this exercise, DALYs are the focus since they are more commonly used in low- and middle-income countries; and
* estimating costs and health outcomes over a sufficient time to capture all important costs and effects for the analysis.

**Interpreting results**

* A lower ICER (change in cost divided by change in effect) means better value for money.
* To establish whether the ICER is cost-effective in each context, a cost–effectiveness threshold is used. The threshold establishes whether money should be used on the intervention, given what is currently being provided in that context.
* Another conclusion from a cost–effectiveness study can be that it is cost-saving, meaning that the intervention provides better health at a lower cost. In this situation, the decision to adopt or maintain the intervention would be very strong based on the cost–effectiveness.
* Cost–effectiveness is usually considered in a specific context, since local prices, disease burden and health system and population health characteristics can influence it.
* Cost–effectiveness estimates can be highly uncertain, given the complexity of factors considered in an estimate. Good studies explore different populations and health systems, scenarios and assumptions and report the levels of uncertainty.

**Scoring cost–effectiveness**

A traffic light system illustrates the cost per DALY averted relative to a country-specific threshold based on the GDP per capita.

|  |  |  |
| --- | --- | --- |
| **Score** | **What it means** | **Criteria** |
| 🟢 **High** | Strong value for money | Consistent evidence suggesting an incremental cost per DALY averted <0.5 times GDP per capita or cost-saving |
| 🟡 **Moderate** | Acceptable value | Consistent evidence suggesting an incremental cost per DALY averted between 0.5 and 1 times GDP per capita |
| 🔴 **Low** | Weak value for money | Limited evidence of an incremental cost per DALY averted >1 times GDP per capita |

**Evidence on cost–effectiveness**

Some interventions are supported by extensive research; others have limited or no data. When evidence is interpreted, thinking in terms of a hierarchy of availability (adapted from GRADE principles) is helpful.

|  |  |  |
| --- | --- | --- |
| **Level** | **Type of evidence** | **Description** |
| I | Systematic review(s) | A high-quality review of multiple cost–effectiveness studies; most reliable |
| II | Multiple cost–effectiveness studies | Consistent findings across studies |
| III | Single cost–effectiveness study | Useful but more limited in generalizability |
| IV | No cost–effectiveness evidence | No published data available |

**How the cost–effectiveness evidence was compiled and categorized for the global exercise**

In this rapid global exercise, there was insufficient time to conduct a full evidence review. The cost–effectiveness data presented here were compiled using available global evidence syntheses (including Disease Control Priorities, Institute for Health Metrics and Evaluation (IHME) and WHO-CHOICE) and literature shared by the WHO Department of Global HIV, Viral Hepatitis and Sexually Transmitted Infections Programmes and targeted PubMed searches intended to address key gaps. This was not a systematic review nor was it intended to serve as a comprehensive or exhaustive source of evidence on cost–effectiveness but rather an indicative starting-point that can help to inform expert judgement and complement expert knowledge on the cost–effectiveness of the interventions.

To facilitate a transparent and cautious interpretation of results, a standard set of inclusion and categorization rules was applied.

* All ICERs were converted to 2023 US dollars using the GDP deflator for the United States of America.
* For each publication, we compared the reported ICER to the GDP per capita of the study setting and applied the categorization thresholds defined earlier (based on 0.5 and 1 times GDP per capita) to classify the results as high, moderate or low.
* Publications were excluded if they met any of the following criteria:
  + no publication year or reference period provided;
  + no clearly identifiable country or region reported (such as unspecified or unclear geographical scope); or
  + the publication date was before 2010.

For publications reporting a range of cost–effectiveness estimates rather than a single point estimate, the following approach was applied.

* For IHME studies, we reported the country-specific median ICERs as presented in the main publication (based on meta-regression estimates) and noted the lower and upper ICER values provided in the supplementary materials.
* For other studies:
  + If both the minimum and maximum ICER values were below the 0.5 or 1 times GDP per capita threshold, the intervention was classified as cost-effective and used the average ICER for reporting purposes.
  + If the ICER range was wide, the intervention was stratified based on key characteristics (such as delivery model or population group). For example, community-based HIV self-testing was disaggregated by target population (young; women having transactional sex; adult men).

**How interventions were summarized**

To synthesize findings across the literature:

* A structured extraction and summary table listing each intervention was created and disaggregated into more specific delivery models or target groups when studies reported results at that level. This was done where the original intervention label was too broad to meaningfully reflect differences in cost–effectiveness across sub-interventions (the intervention as defined in the study).
* For each subintervention:
  + Each ICER was compared with the country-specific GDP per capita threshold.
  + A traffic light score was assigned (high, moderate or low) based on the threshold cut-offs from the table above.
  + The number of ICERs falling into each category was counted.
  + These results were summarized using two pivot tables for visual comparison across interventions, one for evidence only from low- or lower-middle-income countries and one for evidence for all low- and middle-income countries. The table forlow- or lower-middle-income countries should be used for this exercise. The one for low- and middle-income countries is to show the differences by context and ensure that the experts fully understand the difference between evidence that may have been used for developing guidelines but may not apply to this contextual archetype.

These categories do not consider any formal quality grading. Rather, they provide a practical guide for interpretation under time constraints. This evidence process is therefore weak (subject to bias), and experts should treat this classification as indicative only, complementing it with their own knowledge and professional judgement.

**Interpreting gaps, uncertainty and missing evidence**

As such, the strength of the evidence varies considerably across interventions, reflecting context, methods and time point. The evidence should therefore be considered indicative and, in many cases, of weak quality. It is a starting-point for expert judgement and not a definitive assessment. Experts are encouraged to complement this information with their own knowledge and familiarity with the literature, including any additional studies or contextual information they consider relevant.

If no cost–effectiveness evidence is available for a given intervention, this does not mean the intervention should be automatically given lower priority or assumed not to be cost-effective. Depending on how the intervention scores on other priority-setting criteria, it may still be considered for inclusion or flagged for further assessment when more time and resources are available. If the evidence is weak or missing, experts may choose not to assign a score under this criterion.

* + 1. **Budget impact information sheet** **for rapid global HIV, viral hepatitis and STI priority-setting**

**What is it?**

This criterion assesses whether an intervention can be financed and maintained within the country’s health budget, especially when external donor support is uncertain or declining. It assesses the implication for the health budget, now and in the future. In principle, all interventions that are cost-effective should be afforded. However, in some cases when interventions take up large proportions of the budget, there may be concerns about sustaining funding.

Budget impact examines financial resource requirements, now and in the future compared with budget availability (the intervention’s total cost compared with available financial resources). It is concerning that, when giving priority to an intervention requires rapid reallocation, assessing budget impact with high confidence may not be feasible.

It focuses on:

* the cost of the intervention at scale, although periods of scale-up can be included and should be specified;
* the relative size of the resources used compared with available health budgets; and
* in the case of transition from external to domestic spending, impact will depend on whether the intervention is already domestically financed, co-financed or fully donor or externally financed, and the additional cost to domestic budgets may also be considered rather than overall total cost.

**Key concepts**

Budget impact refers to the cost of implementing an intervention (usually at scale) and its relative size compared with the total health budget, which can be estimated by looking at the domestic government health spending (per capita) and the share of spending covered by external aid.

A high budget impact signals that an intervention may:

* crowd out other more cost-effective interventions before full setting of new priorities can be considered; and
* be infeasible in terms of adjusting public financial management systems.

Budget impact analysis is typically conducted from the budget holder's perspective (in this case, the health ministry, finance ministry, social security agency or government health insurance agency) and includes costs and savings but not health outcomes.

**How to assess budget impact**

Assessing budget impact can involve both quantitative indicators and qualitative judgements. It is important to consider total cost (population in need multiplied by unit cost) and not just the unit cost of an intervention. It is also important to think about the full costs: not just commodities but consequences for human resource budgets and even longer-term infrastructure investment and maintenance, for example.

In the global priority-setting exercise, some information is provided below.

* Consider the annualcost of the intervention per capita to annual health budget per capita as a rough benchmark for affordability.
  + Low-income countries domestic budget average (2022): US$ 9 per capita.
  + Low-income countries current external funding average (2022): US$ 12 per capita.
  + Typically, around 60–80% of budgets reach the service delivery levels.
  + Typically, of that 60–70% of the budget, up to half is spent on secondary and tertiary care.

For example, a screening programme may have low cost (less than US$ 1), but a large population (1 million people) may need to be screened, so it still may have a high cost per capita. In a population of 5 million people ((1\*1 000 000)/5 000 000) the annual cost of screening is US$ 0.2 per capita (or 1% of the total health budget for this intervention, with US$ 20 per capita available in the health budget).

**Scoring budget impact**

Use the traffic light categories below to rate each intervention based on its financial implications and likelihood of being sustained through domestic resources.

|  |  |  |
| --- | --- | --- |
| **Score** | **What it means** | **Description** |
| 🟢 **High** | Low budget impact | Low budget impact or easily absorbed into the health sector budget. The average annual cost is very small relative to the health budget or resources are committed to sustain it. Likely to continue without external aid, or alternate financing is secured. |
| 🟡 **Moderate** | Moderate budget impact | Non-trivial cost that poses some strain. May require reallocation of other services currently funded by health sector budgets. Possibly sustainable for a short period or at a reduced scale, but uncertain in the long run without additional funds, efficiency improvements and setting new priorities for health sector budgets. |
| 🔴 **Low** | High budget impact | Very high cost relative to the available health sector budget, making it infeasible to maintain at the current scale, since the budgetary shifts required would be too great. In a scenario of decreased external support, funding this intervention would severely crowd out other services or is simply not possible; continuing it would quickly exhaust resources. |

**Interpreting the evidence on budget impact**

Evidence on budget impact may include:

* budget impact analysis, either in peer-reviewed studies or national policy documents;
* costing studies estimating unit costs and projected scale-up costs;
* programme budgets or expenditure reviews; and
* national plans, Global Fund notes or similar documents.

In this case, no evidence summary is provided since there was insufficient time to estimate the costs per intervention and the reader will need to judge budget impact for a low-income setting (see budget per capita above). In such cases, consider:

* your experience in programme and intervention delivery, including budgeting;
* rapid costing exercises or experiences with which you are familiar ("back of the envelope"); and
* judgements based on similar interventions or country contexts.

Benchmark percentages of spending per capita are not provided as high or moderate and low. For this, reflect on your experience of sustaining programmes or health planning. Similar to the other criteria, you should place this judgement in the context of a low-income country with a generalized epidemic. Similar to the other criteria, this scoring is intended to structure and inform expert deliberations and not to determine the selection of the interventions itself.