Annex 3E. Methods for Appraisal of Essential UHC Interventions

INTRODUCTION

This annex provides details on the methods and data sources used in appraisal of essential UHC (EUHC) interventions described in the accompanying chapter. The appraisal process is meant to facilitate a multi-dimensional assessment of health sector interventions in order to identify a highest priority package (HPP) that is relatively more feasible by 2030 and affordable in the most resource-constrained settings.

The approach of *Disease Control Priorities, 3rd Edition* (*DCP3*) adheres closely to the principles elaborated in the 2014 WHO consultation, *Making Fair Choices on the Path to Universal Health Coverage* (1). Their overall approach was based on principles of fairness and consisted of three steps:

- *Categorize services into priority classes. Relevant criteria include those related to cost-effectiveness, priority to the worse off, and financial risk protection.*

- *First expand coverage for high-priority services to everyone. This includes eliminating out-of-pocket payments while increasing mandatory, progressive prepayment with pooling of funds.*

- *While doing so, ensure that disadvantaged groups are not left behind. These will often include low-income groups and rural populations.*

The first step of this strategy is to categorize interventions into priority classes. The process outlined in this annex mirrors the WHO recommendations. As described in the accompanying chapter, the authors propose four criteria for considering whether to include an intervention in the HPP:

1. Very good value for money in low-income countries
2. Priority given to the worst off
3. Likely to provide a high degree of financial risk protection
4. Part of the "Grand Convergence" agenda proposed by the Lancet Commission on Investing in Health (2).

The rationale for the last criterion is discussed in the chapter and not covered further here. The remainder of this annex describes the approach to grading interventions according to the first three criteria. A separate Annex 3F provides the results of this exercise in these dimensions.

CRITERION 1: VALUE FOR MONEY

As discussed in the accompanying chapter, *DCP3* relies on the use of economic evaluation methods to provide a quantitative assessment of value for money from health interventions. Most commonly, value for money is determined using cost-effectiveness/cost-utility analysis. In the context of low- and middle-income countries (LMICs), the main quantity of interest is usually the incremental cost per disability-adjusted life-year (DALY) averted, compared to a "do nothing" or "standard of care" scenario. Chapter 7 of *DCP3* summarizes and provides critiques of this literature. Where available, geometric means of incremental cost-effectiveness ratios (ICERs) by intervention that were presented in that chapter were applied to EUHC interventions.

Instead of using DALYs as an outcome measure, a number of studies in LMICs have used quality-adjusted life-years (QALYs). The *DCP3* perspective is that the value of a DALY averted and of a QALY gained are roughly similar, and so incremental cost per DALY and cost per QALY ratios are roughly comparable, at least with reference to a willingness-to-pay threshold. For interventions for which ICERs using DALYs were not available, the authors supplemented the database in chapter 7 with literature searches including the Tufts University cost-effectiveness registry (http://healtheconomics.tuftsmedicalcenter.org/cear4/Home.aspx).

A subset of EUHC interventions have only been assessed using cost-effectiveness analysis, with ICERs presented in "natural" units (such as cost per death averted). As described in the chapter, the relevant metric was recorded in these cases and expert judgment applied to determine the relative value for money afforded by these interventions in relation to categories of value for money, described below, that are worked out using DALY-based cutoffs.

DCP3 acknowledges that cost-effectiveness studies are generally accompanied by a high degree of uncertainty and local contextual factors that make quantitative extrapolation to other settings very challenging. (These issues are mentioned in a few places in the accompanying chapter.) To account for the imprecision and uncertainty in generalizing primary data sources, the authors propose five categories of value for money, scored from 4 (best) to 0 (worst):

- 4 = ICER generally less than US$ 250 per DALY averted. This category is the most stringent and reflects the perspective of Claxton and colleagues that the most appropriate willingness-to-pay threshold is based on healthcare opportunity costs, which roughly equates to 0.5 times gross domestic product per capita per DALY averted (3). The reference group for the HPP is low-income countries, where average GDP per capita is currently around US$ 500 according to the most recent World Bank estimates. As stressed in the chapter, DCP3 does not explicitly endorse this cutoff but uses it as an example of the sort of cutoff a low-income country might employ.

- 3 = ICER between US$ 251 and US$ 1300 per DALY averted. The upper limit of this category is 2.3 times the average? GDP per capita of low-income countries. This multiple is based on work done for the *Commission on Investing in Health* that estimated the value of a life-year to be about 2.3 times GDP per capita in LMICs (2); hence interventions with an ICER less than US$ 1300 per DALY averted would be cost-effective (from a value of a life-year perspective) in low-income settings.

- 2 = ICER between US$ 1301 and US$ 4100 per DALY averted. The upper limit is based on the value of a DALY in lower middle-income countries (i.e., 2.3 times GDP per capita in lower-middle-income countries); interventions with an ICER less than US$ 4100 per DALY averted would be cost-beneficial in lower middle-income settings. This cutoff implies that interventions categorized as "2's" or lower are not cost-effective in low-income countries and are therefore much less attractive for the HPP.

- 1 = ICER greater than US$ 4100 per DALY averted. This category captures interventions that are attractive in LMICs generally (especially upper-middle-income countries) but are unlikely to be high-priority for more resource-constrained settings.

- 0 = no economic evaluation data are available for the intervention, nor are there any judgmental reasons to think the intervention is particularly cost-effective.

On the last grouping: the authors acknowledge that a number interventions of recommended in *DCP3* do not have primary estimates of cost-effectiveness but are widely regarded as cost-effective. These interventions were given a score from 1 to 4 based on their overall level of cost-effectiveness according to the judgment of the authors. For a small number of interventions, there was not sufficient or consistent expert judgment available to classify the intervention as anything other than 0. *DCP3* does not recommend *against* such interventions in LMICs but stresses that they would be low-priority in most settings.

Finally, while the scoring system above is simple to implement for DALY and QALY studies, it requires an additional degree of judgment in the case of studies conducted using natural units – including both outcomes that could be converted to DALYs (such as deaths) and outcomes that do not fit into the DALY framework (such as stillbirths, met need for contraception, or suffering relieved by palliative care). These instances are marked as "not applicable" in the cost-effectiveness column of Annex 3F but are graded 0 to 4 on the basis of relative value for money. For instance, highly effective and inexpensive contraceptive methods like oral hormones or implantable devices are graded more highly than more expensive methods such as tubal ligation and vasectomy. As another example, basic palliative care services (centered on the provision of oral morphine for severe life-limiting illness) are graded more highly than expanded palliative care services, which require more human resources for psychosocial support and case management. The latter can be viewed as incremental improvements on (additions to) basic services rather than substitutes.

CRITERION 2: PRIORITY TO THE WORSE OFF

There are a variety of ways in which to identify which individuals in the population are worst off and which interventions most effectively address their health needs. Much of the literature on health equity is concerned with scaling up services to vulnerable groups such as females, children, and poor and rural populations. It is also frequently assumed that major infectious diseases are the most classic examples of "diseases of poverty." Hence, much of the discourse around the Millennium Development Goals (MDG) was focused on control of HIV/AIDS,

tuberculosis, and malaria and on ensuring maternal and child survival, with arguments (appealing to equity principles) that these were the most pro-poor health services.

On the other hand, poor populations suffer from higher disability and mortality rates from a wide variety of causes in addition to those listed previously. The implication of these data is that some other inequality-driven infectious and noncommunicable diseases may be missed by the MDG approach. An alternative perspective on health equity, proposed below, could take a lifecourse approach to health inequality. In this framing, the worst off are those who, by virtue of having a particular disease or injury, have the lowest lifetime level of health. One measure that attempts to reflect this is called health-adjusted age of death (HAAD).

In order to compute HAAD, data on cause-specific incidence, prevalence, average disability, and mortality, as well as all-cause mortality and average population disability, were taken from the Global Burden of Disease 2015 Study (4, 5). (Incidence and prevalence data from this study were used because similar data were not available from the WHO. Mortality data from this study were used in order to remain consistent with the nonfatal measures.)

To calculate condition-specific HAAD, a model cohort of individuals was created for each health condition such that the age pattern reflected the condition-specific incidence. This cohort was then exposed to with-condition mortality risk for a certain length of time, depending on the condition, after which they were exposed to the general population mortality risk. For some conditions, the elevated mortality risk was life-long. Using standard life table methods, the deaths were counted by age. The years lived were then adjusted based on the population average disability in years lived without the condition and on the background disability and condition-specific disability for years lived with the condition.

HAAD, then, is a measure in years of the average health-adjusted age at which individuals with a particular disease or injury die. Aggregate epidemiological and demographic data from low-income countries were used to calculate HAAD, since low-income countries are anticipated to be the primary users of this information in order to develop an HPP.

Interventions were graded on a scale of 1 to 3 on the basis of HAAD cutoffs of less than 40 years (worst, 3) and greater than 50 years (best, 1), with an intermediate category for HAAD 40-49. (The rationale for these HAAD cutoffs is that they reflect realistic decrements in healthy life expectancy in low-income countries; middle-income countries may choose higher cutoffs since their populations tend to be older and with higher healthy life expectancy.)

HAAD for each intervention was taken from the most-specific Global Burden of Disease cause that the intervention addressed. Interventions that addressed multiple causes were assigned to higher-level causes that encompasses all causes addressed. For example, "Long term management of ischemic heart disease, stroke, and peripheral vascular disease with aspirin, beta blockers, ACE-inhibitors, and statins (as indicated) to reduce risk of further events" was assigned to "cardiovascular diseases" (level 2) rather than "ischemic heart disease," "ischemic stroke," or "peripheral vascular disease" (levels 3-4) since the intervention has an impact on all three conditions.

CRITERION 3: FINANCIAL RISK PROTECTION

Quantitative estimates of financial risk protection (FRP) that are comparable across interventions are challenging for several reasons. First, to date no one modeling approach has been developed that can encompass all of the interventions in *DCP3* and calculate reductions in hard outcomes (such as poverty cases averted) in a consistent manner. (An extended cost-effectiveness analysis in Ethiopia, conducted alongside *DCP3*, provided a comparative assessment of nine interventions, but this work has not yet been extended or applied to a wide variety of interventions or in other country settings (6, 7).) Second, which financial outcome (or outcomes) is (or are) the most relevant and externally valid measure(s) of financial risk is an as-yet unsettled, empirical question. Third, acute and chronic illnesses may have different patterns of financial risk (and consumption smoothing) in the long run, making it challenging to identify a relevant time horizon for an analysis across disease groups.

Nevertheless, FRP is an important outcome for universal health coverage schemes and cannot be ignored in the priority-setting process. To provide some sense of relative FRP afforded by various interventions in DCP3, the authors developed a composite indicator of FRP using existing data. (Future work will seek to validate this measure empirically; it remains illustrative for the purposes of this chapter.) According to the *Making Fair Choices* report:

Several factors increase the amount of direct protection associated with coverage of a particular service. These factors include high cost and out-of-pocket payments, low predictability of need, pronounced urgency and severity of the target condition, and high incidence of the target condition among the poor.

If these factors are assumed to have face validity, a scoring system can be constructed based on the properties of each intervention and the cause(s) addressed:

A. Is the intervention, by virtue of its high cost, likely to result in medical impoverishment in the absence of public finance?

A crude quantitative measure of potential impoverishment is the size of the unit cost of the intervention. All else being equal, interventions that have higher copayments are more likely to lead to catastrophic/impoverishing expenditure, hardship financing, or other economic consequences. The distribution of unit costs from EUHC interventions is shown below. If current GNI per capita for low-income countries is US$ 830 (in 2012 USD), then daily income in this group is about $2.3, weekly income is about US$ 16, and monthly income is about $69. About 25% of these unit costs are less than daily income, 60% less than weekly income, and 73% less than monthly income.
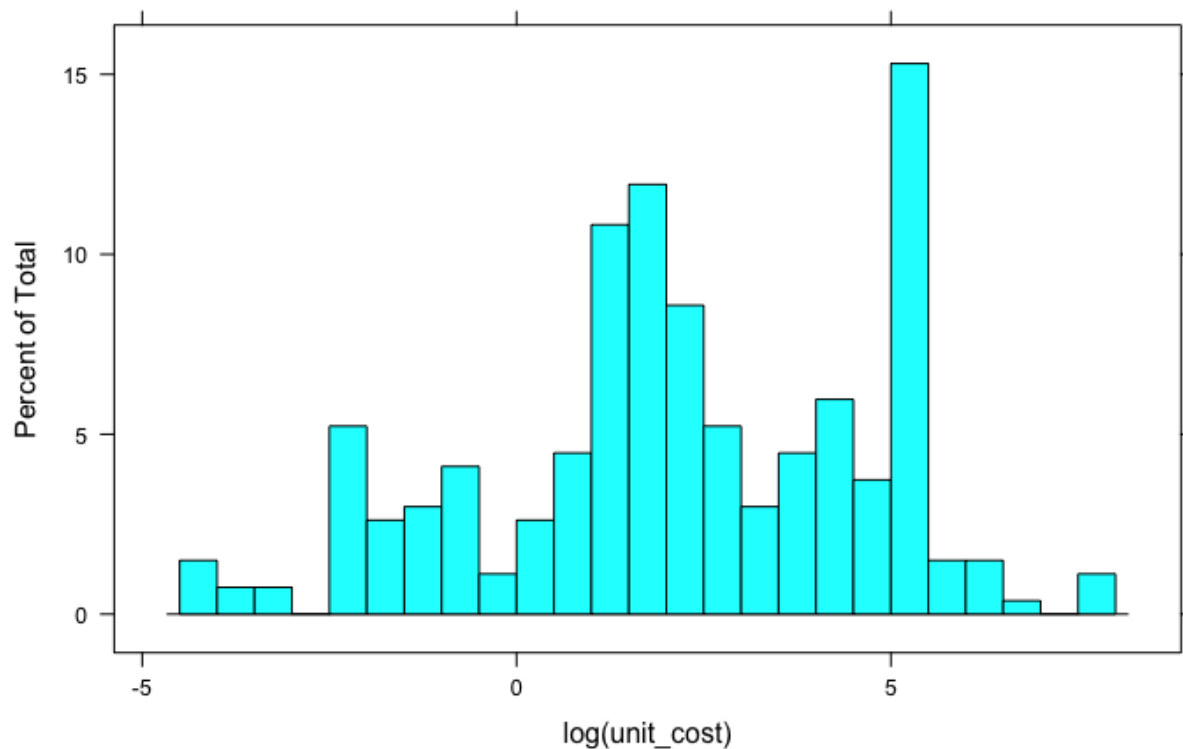
Figure 3E.1 Distribution of unit costs in this chapter (log scale)

One could argue that unit costs exceeding average daily income in an average low-income country would put an individual at mild financial risk, weekly income at moderate financial risk, and monthly income at high financial risk. A zero-to-three-plus system could hence be implemented for this indicator on this basis. (For the purpose of this exercise, the scoring was done for unit costs in low-income countries, the focus of the HPP. For lower middle-income countries, the proportion of costs in these groups would roughly similar because the gradients in unit costs are generally proportional to country income – i.e., that most aspects of healthcare are non-tradable).

B. Does the intervention address an unpredictable acute event?

This criterion draws on the prior assessment of intervention urgency that was developed for each *DCP3* essential package. Interventions that are acute and are delivered in hospitals would score "1" for this indicator.

C. Does the condition addressed by the intervention cause high rates of disability or death among working-age populations?

For this criterion, the primary cause addressed by the intervention was assessed as for HAAD above. Two metrics were computed: average disability weight (defined as total YLDs divided by total prevalent or incident cases) and average loss of life expectancy (defined as YLL/death).

Interventions that address a cause whose disability weight exceeds 0.1 or for which the average loss of life expectancy exceeds 46 years (i.e., where the average age of death is 86 [global reference for YLL] – 46 = 40 years, a reasonable midpoint for working age adults) would score "1" on this measure (or a "2" if the condition meets both thresholds).

A final indicator was added to these three: whether or not the intervention would be likely to crowd out out-of-pocket expenses due to less effective, more expensive therapies that tend to impoverish households in low- and middle-income countries. Palliative care services are a good example of this: by increasing access to interventions that relieve suffering, households will be less likely to choose therapies – such as chemotherapeutics (in the case of many advanced solid tumors), or traditional medicines with no known medical benefits – that provide little survival or quality of life benefit and are very costly. Interventions that met this criterion were assigned a score of "1."

The overall FRP score, then would be the sum these indicators, with possible ranges from 0 to 7, similar to what has been defined for the equity and cost-effectiveness rankings in this chapter. This scoring system reflects all the factors listed in the *Making Fair Choices* report with the exception of the "high incidence of the target condition among the poor" factor. However, the authors argue that the construction of HAAD and emphasis on severe diseases in this epidemiological context (i.e., low-income countries) adequately deals with "the poor" from a global perspective.

REFERENCES

1. WHO. Making fair choices on the path to universal health coverage. Final report of the WHO Consultative Group on Equity and Universal Health Coverage. Geneva: World Health Organization; 2014.
2. Jamison DT, Summers LH, Alleyne G, Arrow KJ, Berkley S, Binagwaho A, et al. Global health 2035: a world converging within a generation. Lancet. 2013;382(9908):1898-955.
3. Ochalek J, Lomas J, Claxton K. Cost per DALY averted thresholds for low- and middle-income countries: evidence from cross country data. Centre for Health Economics (CHE) Research Paper 122. York: University of York; 2015.
4. GBD 2016 Mortality Collaborators. Global, regional, and national under-5 mortality, adult mortality, age-specific mortality, and life expectancy, 1970-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet. 2017:in press.
5. GBD 2015 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 310 acute and chronic diseases and injuries, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet. 2016;388:1545-602.
6. Verguet S, Olson ZD, Babigumira JB, Desalegn D, Johansson KA, Kruk ME, et al. Health gains and financial risk protection afforded by public financing of selected interventions in ethiopia: An extended cost-effectiveness analysis. The Lancet Global Health. 2015;3(5):e288-e96.
7. Verguet S, Memirie ST, Norheim OF. Assessing the burden of medical impoverishment by cause: a systematic breakdown by disease in Ethiopia. BMC Med. 2016;14(1):164.