

Appendix 2: Methodology

Conceptual Model

The dataset covered in this report includes data for health R&D efforts of high-income countries, low- and middle-income countries, and an area of convergence of health R&D funded by HIC and carried out in and for the primary benefit of LMIC. The latter includes research funds transferred from high-income countries to low- and middle-income countries, health research carried out in LMIC by HIC entities, R&D carried out in high-income countries that is relevant to the needs of low- and middle-income countries and health research carried out in high-income countries by researchers from low- and middle-income countries.

Sources of data

OECD database

The OECD database is designed to provide analysts with comprehensive and internationally comparable time series on R&D (including health-related R&D) by sectors of performance and sources of funds when such data are available. The OECD database has been created for 19 of the largest R&D performing countries, and also provides a zone total for the European Union. LMIC that publish R&D expenditures are usually the ones with major contributions to global health R&D. The subset of LMIC documented may not be representative of other low-income countries.

ANBERD database

The ANBERD database (Analytical Business Enterprise R&D) was developed with the objective of creating a consistent data set that overcomes the problem of international comparability and provides for time series consistency. In conformity to the Frascati definitions and methodologies, ANBERD incorporates annual R&D expenditure carried out in the business enterprise sector, regardless of the origin of funds, for a variety of industrial sectors. Several factors may have a substantial effect on the consistency of ANBERD, such as the completeness of the set of enterprises for which R&D data are collected and classification scheme for firms associated with more than one sector. It is difficult to maintain consistency over time due to these factors, hence time series may have large discontinuities.

RICYT database

The RICYT (Red Iberoamericana de Indicadores de Ciencia y Tecnologia) database provides both a percentage distribution of S&T and/or R&D expenditure by socio-economic objectives as well as total national R&D time series for several countries in South and Latin America, Spain, Portugal and the USA.

Data for public sector expenditures for international health research

Data on health R&D funded by Official Development Assistance (ODA), bilateral and multilateral agencies, development banks and national health research institutes were also

compiled by consultants under contract to the Global Forum. Data sources include material obtained from various official organizational publications and web sites, and through personal communications to Global Forum consultants from organizational officials.

Private not-for-profit sector data

Funding data for health research by foundations and other not-for-profit organizations were compiled by consultants under contract to the Global Forum. Data sources include The Foundation Center for U.S. Foundations and the European Foundation Center (EFC), material obtained from various official organizational publications and web sites, and through personal communications to Global Forum consultants from organizational officials.

Private sector data

Data for the private sector were obtained from the performer base of the OECD and RICYT databases and from data from official publications and web sites of PhRMA and other pharmaceutical associations, and other publications detailing private sector expenditures.

Other data sources

Various other data sources were used to fill in, supplement and contextualize data from the above sources. These include publications from UNDP and WHO for data on public health expenditures; WHO for data on injuries and communicable diseases; the World Bank Group for GDP data; the World Bank for data on purchasing power parity and absolute poverty; and data from web sites, publications and communications from organizations such as the Kaiser Foundation and various international initiatives, such as the International Aids Vaccine Initiative (IAVI) and the National Institute of Allergy and Infectious Diseases (NIAID), for disease-specific research.

Specific references to these and other data sources are cited throughout the text, charts and tables in the report as appropriate.

Estimations and Methodologies

Performer reported R&D

The reference manual for methodology regarding the monitoring of R&D expenditures used by the OECD is the Frascati Manual, and the Bogotá manual for RICYT data. The maximum of R&D by health objectives and fields of science was used for health research with the OECD data. Note that R&D by health objectives and fields of science are generally consistent with each other. For the RICYT data, a portion of advancement of knowledge (research in higher education) was included as health research.

Pharmaceutical R&D

Pharmaceutical data in this report include data covering preclinical trials, Phase I to IV of clinical trials, drug approval and marketing.

Health R&D updates

For several countries such as India, South Africa, Australia and others, health R&D data were collected directly from national surveys, pharmaceutical associations and other publications. In such instances, the OECD, ANBERD and RICYT figures were updated accordingly.

Methods of estimation for missing data

Various factors resulted in incomplete or missing data for health R&D expenditures for a given country and year:

- Missing health R&D expenditure for the government or business enterprise sector.
- Unknown distribution of health R&D expenditure of a sector by sources of funds.
- Unpublished health R&D expenditures for the private not-for-profit sector.
- Missing data-point on a time series. Most countries report/publish health R&D data at regular time intervals, such as odd or even years, although some do so at irregular time intervals.

In such instances, methods of estimation have been used when appropriate to estimate the missing health R&D expenditures. Whenever possible, the health R&D expenditure for a missing year was extrapolated from the trend in GERD. The interpolation method was used by default, provided the outer data points were at reasonable time proximity from the missing point. This method is practical and easy to apply, and the error is limited to the extent that the two external points are accurate. Extrapolation methods not based on a secondary trend indicator such as GERD were avoided as the error resulting from such estimates is generally unbounded. The health R&D expenditure for a given sector was redistributed by sources of funds, such as distributing according to a preceding year's figures, distributed according to total the GERD, based on rules.

Reconciliation between 2001 health research and 1998 reported figures

A retroactive adjustment of the previously reported 1998 figures was used for consistency purpose. The 2001 figure for the global health R&D expenditure is US\$105.9 billion versus US\$73.5 billion published in 1998. This increase resulted from the growth in health R&D sources of data.

The adjustment due to changes in the estimation methodology was computed from the actual growth in total health R&D expenditure from 1998 to 2001, using time series of R&D expenditures from OECD. Due to the limited set of historical data available in LMIC, the corresponding growths may not be representative of all LMIC. Variation in local currency exchange rate relative to US dollars may have a substantial effect on the growth, as health research expenditures were converted to current US dollars for each country.

Global Burden of Diseases (GBD)

WHO has undertaken a new assessment of GBD for 2002 and this is used to provide an overview of the main causes of burden of disease in low- and middle-income countries and of major trends since 1990. The data sources and methods used in the GBD2000 study are documented on the World Wide Web (www.who.int/evidence/bod).

Principal risk factors of inaccuracies in estimations

The actual flows of funding for health research are difficult to trace due to various factors such as incompleteness, misclassification, risk of double counting, incorrect recognition of expenditures, and inconsistency between sources of data. Each of these may have varying and potentially substantial effects on the estimates.

Completeness/incompleteness

Health research expenditure data were not available for all countries. Indeed, performer reported data were available for most HIC but not for all LMIC. OECD provided performer reported data for LMIC that likely account for the greatest proportion of health research expenditures among these countries.

Risk of misclassification

Classification may have a substantial effect on the estimates of health R&D expenditures and their consistency over time. Changes in methods of classification vary, resulting in variations from one reporting period to the next, or from one country to another. For example, firms may be associated with more than one sector making it difficult to maintain a consistent classification scheme.

Risk of double counting

There are two sources of risk leading to double counting of health research flows.

(1) Flows through multiple agencies

Whenever funding flows through several agencies or other entities, a risk of double counting arises. The first entity may fund a project through a second entity and the amount is reported twice. Hence, the whole flow of funding needs to be tracked from source to final point of expenditure.

(2) Intramural and extramural expenditures

Due to the nature of the health research sector, the trend in the business enterprise sector is to outsource some research to other laboratories, research institutes or hospitals; for instance, to conduct clinical studies. This may arise, among other considerations, from the decision to focus on the value-added activities of the firm and the capacity to invest in expensive equipments, consumables and the maintenance of a laboratory. Although most large pharmaceutical companies own some research facilities, access to such services is critical to the development of smaller pharmaceutical companies.

To avoid double counting, extramural expenditure has been subtracted from the pharmaceutical R&D whenever possible. Unfortunately, the extramural pharmaceutical R&D expenditure often is not identified.

(3) Health research ODA and public funding in LMIC

Some of the ODA health research expenditures in LMIC may be included in the government sector. Hence, there is a risk of double counting between health research ODA and public expenditures on health in countries receiving ODA.

Recognition of expenditures

Health research expenditure estimates are based either on performers' surveys or from funders directly. Some variability may arise because such surveys may not include all funders or performers, although typically they do account for all the major players. Performer-based expenditures, either budgeted or actual, are recognized in the year expenses are incurred. In contrast, transferred funds from the funder may not be expended by the recipient during that same year, but rather over several years.

Consistency across different sources of data

Financial data on health research have been gathered from a variety of sources such as the OECD, RICYT and ODA data, The Foundation Center, pharmaceutical associations, national surveys and publications among others. The different classifications used for health R&D make it difficult to assure consistency across different sources of data. Estimates were produced using consistent assumptions and methodologies to assure the best possible data.