ONLINE ANNEXE A

Additional methodological considerations

IDENTIFICATION OF SURVEY RECIPIENTS

Year one G-FINDER survey recipients were identified through various avenues including our own contacts database; previous neglected disease surveys in HIV/AIDS, tuberculosis (TB), and malaria; and research to find previously unknown funding organisations in countries with high research and development (R&D) expenditure per gross domestic product (GDP).

In 2008, we focused on groups and countries that were missing or poorly represented in year one, developing proactive strategies to both increase the number of survey recipients and improve response rates in these areas. Major Indian public agencies involved in funding R&D for neglected diseases were identified and incorporated in our list of participants, and additional diagnostics organisations and small pharmaceutical and biotechnology firms (SMEs) were also included. In 2009, the survey was expanded to capture major public funding agencies in an additional three developing countries (DCs), Ghana, Colombia and Thailand, and in 2010 expanded again to reach public funders in Argentina, Chile, Malaysia, Mexico, Nigeria and Uganda. In 2011, several organisations known to be active in malaria R&D were surveyed for the first time as part of a project to measure R&D funding into malaria elimination- and eradication-specific activities conducted on behalf of the Malaria Eradication Scientific Alliance (MESA). Since then, any time that new diseases have been added to the survey scope, organisations known to be active in these areas have been identified and surveyed.

EXPANSION OF G-FINDER SCOPE

In 2013, following consultation with the new international Advisory Committee (AC) (see Annexe 2), the survey was expanded in accordance with our established neglected disease criteria (the disease disproportionately affects people in DCs; there is a need for new products; and there is market failure) to include three additional diseases: cryptococcal meningitis, hepatitis C genotype 4 and leptospirosis. The AC also identified the emergence of a significant potential commercial market for dengue vaccines. As a result, dengue vaccine R&D was removed from the scope of the survey, and data on these investments excluded retrospectively. Other dengue products continue to be included.

In response to the 2014 West African Ebola epidemic, the survey scope was expanded again in 2014 to capture investments in Ebola R&D for diagnostics, drugs and preventive vaccines, as well as basic research. On the advice of the AC, the scope of the hepatitis C category was also expanded in 2014 to capture investment into R&D for two additional genotypes (5 and 6) that disproportionately affect people in DCs.

In 2015, the survey introduced the new grouped disease category of African viral haemorrhagic fevers (VHFs). In addition to Ebola, which was already part of the survey, this new category allowed respondents to report R&D funding for Marburg and Other and/or multiple African VHFs. All funding in this category (including 2014 Ebola data) has been analysed separately to the core G-FINDER neglected diseases in order to avoid the distorting effect this funding had on the analysis. The scope for *Streptococcus Pneumonia* vaccines was also revised to better reflect current approaches to developing pneumococcal vaccines for low-resource settings.

RESTRICTIONS ON SPECIFIC DISEASE-PRODUCT AREAS

Following the methodology used in previous years of the G-FINDER survey, only investments specifically targeted at DC needs were eligible for inclusion in R&D areas where commercial overlap was significant. For instance, a vaccine for *N. meningitidis* should provide coverage against *N. meningitidis* serotype A, be a conjugate rather than a polysaccharide vaccine, be designed for use in infants less than two years of age and be designed to cost less than a dollar per dose. (See Table 1 in the main report for full inclusions for G-FINDER, and the G-FINDER 2008 report for a full description of the original methodology to identify DC-specific investment.)

HANDLING OF FINANCIAL DATA

The following key financial data collection principles were used:

- Survey recipients were asked to enter grant-by-grant expenditures incurred during their financial year (as opposed to the 2015 calendar year) that had the largest overlap with 2015. Intermediaries and product developers were also asked to enter grant-by-grant revenue during the same period
- Only expenditures were included, as opposed to commitments made but not yet disbursed or 'soft' figures such as in-kind contributions, costs of capital, or funding estimates
- All figures in the year nine G-FINDER report have been adjusted for inflation and are reported in 2015 US dollars (US$). All reported data was adjusted for inflation using consumer price index (CPI) estimates from the International Monetary Fund (IMF)[115] and any data entered by survey participants in their local currency was converted to US$ based on the 2015 average annual exchange rate as reported by the IMF,[116] Bank of England,[117] United Nations Treasury[118] and OANDA.[119]

SURVEY TOOL AND PROCESS

As in previous years, the following core principles guided the G-FINDER survey:

1. Only primary data reported by the funders, intermediaries and product developers themselves were included in the survey. No secondary data or estimates were included
2. All primary grant data was collected using the same online/offline reporting tool and inclusion/exclusion framework for all survey recipients.

The main exception to the second principle above was the US National Institutes of Health (NIH), where grants were collected using the Research Portfolio Online Reporting Tools (RePORTER) and the Research, Condition and Disease Categorization (RCDC) process. The information mined from these publicly available databases was then supplemented and cross-referenced with information received from the Office of AIDS Research (OAR) and the National Institute of Allergy and Infectious Diseases (NIAID). Funding from the Innovative Medicines Initiative (IMI) was retrieved from their projects website.

Survey tool

Following the methodology used in previous years of G-FINDER, survey participants were asked to enter every neglected disease investment they had disbursed or received in their financial year 2015 into a password-protected online database, including the grant amount, grant identification number, a brief description of the grant and the name of the funder or recipient of the grant. New survey recipients were also asked to confirm their organisation details such as role in funding (e.g. funder, fund manager, product developer), financial year, currency used, type of organisation (e.g. private sector firm, academic institution, multilateral organisation), and country where they were located. Each grant was entered using a three-step process where the survey recipient had to choose (1) a specific disease or sub-disease; (2) a product type (e.g. drugs, vaccines, microbicides); and (3) a research activity within the product type (e.g. discovery and preclinical, clinical development); according to pre-determined categories as described in Table 1 in the main report. Where survey recipients could not provide data to this level of detail, they were asked to provide the finest level of granularity they could. If survey recipients were not able to allocate the grant to a single disease in step 1, three options were available:

- 'Core funding of a multi-disease organisation' (e.g. funding to an organisation working in multiple diseases, where the expenditure per disease was not known to the funder)
- 'Platform technologies', further allocated as investment into diagnostic platforms; adjuvants and immunomodulators; or delivery device platforms. These categories aimed to capture investments into technologies which were not yet directed towards a specific disease or product

- 'Unspecified R&D' for any grants that still could not be allocated.

Data sharing with other surveys

Primary grant data for HIV/AIDS was shared with and between the HIV Vaccines and Microbicides Resource Tracking Working Group to avoid re-surveying funders when possible. Any primary grant data received by other groups was reviewed and reclassified according to G-FINDER guidelines prior to entry into the database.

DATA CLEANING

Survey closure was followed by a two-month period of intensive cleaning, cross-checking, and organising of the complex dataset collected. This followed a three-step process:

1. All 9,070 grants were reviewed against our inclusion criteria and checked for correct allocation to disease, product type and research type

2. Automated reconciliation reports were used to cross-check 'disbursed' funding reported by funders against 'received' funding reported by recipients (i.e. intermediaries and product developers)

3. Discrepancies were solved through direct contact with the funder and recipient to identify the correct figure. The threshold for discrepancy checking was $0.5m (i.e. any grant over 0.01% of total funding), with the exception of a few major funders, in particular the US NIH, where the threshold was increased to $2.0m. In the few cases where discrepancies still remained, the funder's figures were used.

Industry figures were reviewed against industry portfolio information held by Policy Cures Research and against Full-Time Equivalent (FTE) and direct costs provided by other companies. Costs that fell outside the expected range, for example, above average FTE costs for clinical staff, were queried and corrected with the company.

LIMITATIONS TO INTERPRETATION

Potential limitations with any survey, including G-FINDER, are:

Survey non-completion

A total of 185 organisations participated directly in the G-FINDER survey, reporting data on behalf of a total of 209 organisations. This meant that we received data for more organisations than the previous year, despite targeting our survey follow-up to increase efficiency.

Some neglected disease R&D funding might not have been captured because organisations were not identified as active in this field and therefore were not invited to participate, or were invited to participate, but did not respond. Despite this, we are confident that the majority of neglected disease R&D funding has been captured by G-FINDER because large funders active in this area and target groups identified by the AC were prioritised on follow-up.

Time lags in the funding process

Time lags exist between disbursement and receipt of funding as well as between receipt of funds and the moment they are actually spent. Thus, grants by funders will not always be recorded as received by recipients in the same financial year and there may be a delay between R&D investments as reported by G-FINDER and actual expenditure on R&D programmes by product developers and researchers. Nevertheless, as this report analyses trends over a nine-year period, the impact of time lags is minimal.

Inability to disaggregate investments

Funding allocated to some diseases and products may be slightly underestimated due to:

- Multi-disease organisations: Core funding grants to organisations working on multiple diseases, such as the European & Developing Countries Clinical Trials Partnership, are not counted within the funding figures for specific diseases

- Multi-disease grants: When funders were unable to disaggregate multi-disease grants, these investments were included in the 'Unspecified R&D' category. This is likely to particularly affect US NIH figures for individual diseases. This methodology was followed to prevent double counting investments from the US NIH and is also the reason why the G-FINDER figures do not match the RCDC figures (e.g. categories used in the RCDC system are not mutually exclusive and multi-disease grants are reported fully under all relevant diseases, with risk of double counting).

Non-comparable data

Due to a significant increase in the size of the survey in 2008, data from 2007 is the least comparable to other years. To avoid reporting on artefactual changes related to survey participation, this report only highlights increases or decreases reported by repeat survey participants (YOY funders), which represent real funding changes. Furthermore, the current public official databases for the US NIH data, the RCDC and RePORTER, used for data collection between 2008 and 2015, use a different structure than the US NIH database used in 2007. This means reports obtained from RCDC and RePORTER in years two to nine are not directly comparable to those used in year one.

Missing data

G-FINDER can only report the data as it is given to us. Although strenuous efforts were made to check the classification, accuracy and completeness of grants, in a survey this size it is likely that some data will still have been incorrectly entered or that funders may have accidentally omitted some grants. We believe, however, that the checks and balances built into the G-FINDER process mean that such mistakes, if present, will have a minor overall impact.

VARIATION BETWEEN SURVEYS

Other groups also publish annual surveys of global R&D investment into selected neglected diseases, such as HIV/AIDS and TB. Although G-FINDER worked in close collaboration with some of these groups, both to ease survey fatigue on the part of participants and to clarify any major variance in our findings, each survey nevertheless has slightly different figures. This is chiefly due to differences in scope, in particular inclusion in other surveys of funding for advocacy, capacity building and operational studies – all excluded from G-FINDER. Methodological differences also lead to variations, in particular that G-FINDER figures are adjusted for inflation and exchange rates, which is not always the case for other surveys. As mentioned above, classification of some funding as 'unspecified' in G-FINDER (e.g. multi-disease programmes) may in some cases also lead to different figures than those published in disease-specific surveys.