TECHNICAL REPORT

Current and future burden of communicable diseases in the European Union and EEA/EFTA countries – Methodology protocol

www.ecdc.europa.eu

ECDC TECHNICAL REPORT

Current and future Burden of Communicable Diseases in the European Union and EEA/EFTA countries (BCoDE)

Methodology protocol

National Institute for Public Health
and the Environment
Ministry of Health, Welfare and Sport

This protocol was commissioned by ECDC and coordinated by Piotr Kramarz, Andrew Amato and Alessandro Cassini. It was produced by Marie-Josée Mangen, Cheryl Gibbons and Mirjam Kretzschmar on behalf of the BCoDE Consortium[†].

[†]Dr. Mirjam Kretzschmar (project leader), Dr. Ardine de Wit (RIVM – National Institute of Public Health and Environment, Bilthoven, the Netherlands/UMC – University Medical Center, Utrecht, Netherlands); Prof. Dr. Arie Havelaar (RIVM/University of Utrecht); Alies van Lier, Paul Bijkerk (RIVM), Dr. Marie-Josée Mangen (UMC Utrecht); Prof. Dr. Uwe Siebert, Dr. Beate Jahn, Dr. Nikolai Mühlberger (UMIT – University for Health Sciences, Medical Informatics and Technology, Hall i.T., Austria); Prof. Dr. Alexander Krämer, Dietrich Plaß, Dr. Paulo Pinheiro (University of Bielefeld, School of Public Health, Bielefeld, Germany); Prof. Dr. Walter Ricciardi, Prof. Dr. Elisabetta Franco, Dr. Silvia Longhi, Laura Murianni, Chiara de Waure (UCSC – Università Cattolica del Sacro Cuore, Rome, Italy); Dr. Eric Fevre, Cheryl Gibbons (University of Edinburgh, United Kingdom); Dr. Taavi Lai (Ministry of Social Affairs, Tallinn, Estonia); Dr. Kristi Rüütel (National Institute for Health Development, Tallinn, Estonia).

The BCoDE consortium would like to thank Arun Nanda (WHO Regional Office for Europe, Denmark), Karl Ekdahl and Massimo Ciotti (ECDC), Claudia Stein (WHO, Geneva), Theo Vos (University of Queensland, Australia), Sergio Mariotti (ISS – Istituto Superiore di Sanità, Italy), Philippe Beutels (University of Antwerp, Belgium), Carline van den Dool and John Brooke (UMC Utrecht, the Netherlands), Peter Achterberg and Roel Coutinho (RIVM, the Netherlands), Dr. Jevgenia Epstein (Health Board, Tallinn, Estonia); Ruth Schwarzer, Petra Schnell-Inderst and Johannes Wurm (UMIT, Hall i.T., Austria), Caterina Rizzo (ISS, Italy) and Silvio Capizzi (UCSC, Italy) for critical feedback during the development of this methodology protocol.

Suggested citation: European Centre for Disease Prevention and Control. Methodology protocol for estimating burden of communicable diseases. Stockholm: ECDC; 2010.

Cover picture: © Flickr/Phil Photostream

Stockholm, June 2011

ISBN 978-92-9193-294-8

doi 10.2900/54348

Table of contents

Abbreviations

BCoDE	Current and future Burden of Communicable Diseases in the European Union and EEA/EFTA countries
BoD	Burden of Disease
DALY	Disability-Adjusted Life Year
DSP	Disease-Specific Programme, ECDC
GBD	Global Burden of Disease study, WHO
PCP	Primary care physician
PTO	Person Trade-Off
QALY	Quality-Adjusted Life Year
SEYLL	Standard Expected Years of Life Lost
SG	Standard Gamble
TTO	Time Trade-Off
VAS	Visual Analogue Scale
YLD	Years Lived with Disability (or Years Lost due to Disability)
YLL	Years of Life Lost

Executive summary

Background

The European Centre for Disease Prevention and Control (ECDC) has a responsibility to identify, assess and communicate current and emerging threats to human health from communicable diseases. Surveillance data follow trends in time and indicate health problems resulting from particular diseases, but are subject to under-reporting and do not account for disease severity. To increase insight into the impact of communicable diseases on population health in Europe and to support health policy-making, in 2008 ECDC initiated a call to develop a methodology for measuring the current and future burden of communicable diseases in the European Union Member States and European Economic Area (EEA)/European Free Trade Association (EFTA) countries. The result was the current and future Burden of Communicable Disease (BCoDE) project.

The current protocol is the output of Work Package 1 of the BCoDE project. Choices concerning the methodology of future disease burden estimates were taken on the basis of literature reviews, data reviews and the results of workshops/expert meetings. These are summarised in the current protocol.

Objectives

The primary objective of the BCoDE project is to generate evidence-based, valid and comparable estimates of the burden of communicable diseases and related conditions on society in EU Member States and EEA/EFTA countries. The aim of Work Package 1 was to assess and define the methodology used to generate these estimates. In defining the tools for the generation of estimates one rationale was to ensure comparability with earlier burden estimates published by WHO.

Methods

The communicable diseases for which disease burden will be calculated were selected on the basis of a consensus reached by the ad hoc Advisory Forum working group, resulting in a predefined list of inclusion/exclusion criteria. The disease burden of the selected diseases will be calculated for all Member States of the European Union and EEA/EFTA countries. The disease burden will be estimated using DALYs (Disability-Adjusted Life Year), a health gap measure representing years of life lost (YLL) due to premature death and number of life years lost due to disability (YLD). To better represent communicable diseases and their associated sequelae, we will use the pathogen-based DALY approach which links sequelae to their cause of infection (pathogen). Disease burden caused by sequelae is thereby attributed to incident infections. This can be done by taking an incidence and pathogen-based approach, where the incidence of infections is linked to all possible sequelae through an outcome tree or disease progression model. The outcome trees used in this study are based on literature reviews and have been analysed by ECDC's Disease-Specific Programmes (DSP). Health states are considered part of the outcome tree if there is evidence of a causal relationship between infection and the health state.

In the disease burden calculations YLL will be estimated using the Standard Expected Years of Life Lost (SEYLL) based on the highest observed life expectancy, which is that of the Japanese population. A European life expectancy will be used as an alternative. Severity weights (i.e. disability weights) for non-fatal health outcomes were obtained from the WHO Global Burden of Disease (GBD) study and will be updated when new weights become available. In conditions for which no weights exist, weights will be adapted from existing GBD severity weights for similar conditions. Time discounting and age-weighting will not be applied in the base case analysis, but it will be possible to do so if necessary. Reported incidences will be averaged over a three-year period (2005-2007) to ensure the stability of estimates. The resulting averages will be checked for representativeness in relation to longer reporting periods and, if necessary, the time span for the averaging will be pathogen-based. Data on mortality will be used from both national sources and EUROSTAT with correction for under-reporting where necessary. Attributable fractions and modelling will be used for fatal cases of disease with multiple causes.

Surveillance data, hospitalisation data, primary care data and disease-specific registration databases will be used to estimate incidences of the various non-fatal health outcomes. However, these data sources may only represent part of the true incidence due to under-reporting and under-ascertainment. Therefore, methods will be applied to correct for under-reporting. Attributable fractions and modelling will also be used for non-fatal health states with multiple causes. The duration of health states will be based on national estimates, where available. Disease burden models will be implemented in Excel using @Risk (an add-in to Excel) in order to explore uncertainty. The outcomes of the model are disease burden estimates associated with a specific infectious

disease and its related sequelae in a particular country. The estimates will be presented as DALYs/YLL/YLD per year per country, DALYs per 100 000 population per year and DALYs per year per infected case. Point estimates (means) of these quantities, together with measurements of their uncertainty (e.g. 95% credible intervals) will be presented and discussed and, where necessary, scenario analysis conducted. The current modelling approach assumes steady-state and is therefore not suitable for predicting disease burden into the future.

Perspective

Using the above approaches it is possible to obtain disease burden estimates that enable comparison among communicable diseases and an assessment of their relative impact on public health. Furthermore, the estimates obtained enable comparison with other factors affecting public health, both within and among countries. Thus, the burden of disease estimates will make it possible to prioritise future monitoring and intervention efforts for communicable diseases. The methodology described in this protocol represents a first step in the development of methods to estimate disease burdens for communicable diseases. Future work will extend and improve the methods described by taking temporal dynamics and demographic developments into account. Therefore this protocol should be seen as a working document that will be subject to change as the project progresses in future work packages.

Purpose

Why the BCoDE project?

The European Centre for Disease Prevention and Control (ECDC) has a responsibility to identify, assess and communicate current and emerging threats to human health from communicable diseases.[1] As part of its efforts to meet this responsibility, ECDC produces the *Annual Epidemiological Report on Communicable Diseases in Europe.*[2] The data reported describes trends and, although it can be used to indicate health problems due to particular diseases, it is not suitable for a comprehensive assessment of disease burden. In order to make such an assessment information is required on the severity of disease, including morbidity due to sequelae and disease-related mortality.

In order to meet the above-mentioned responsibility, in 2008 ECDC initiated a request for proposals to develop a methodology to measure the current and future burden of communicable diseases in the EU Member States and EEA/EFTA countries. This resulted in the Burden of Communicable Disease in Europe (BCoDE) project.

The current methodology protocol was developed in Work Package 1 and describes the methodology to be applied when measuring the burden of communicable disease in EU Member States and EEA/EFTA countries under the BCoDE project.

Steps towards the methodology protocol

A metric for communicable diseases

Various metrics exist for capturing and weighing the specific symptoms, severities, chronic sequelae and incidence of morbidity and mortality associated with each hazard as one single integrated unit. Both monetary and non-monetary integrated measures may be applied to value health effects. Given the numerous method options available, an investigation was necessary to find the one most suitable for describing burden of communicable disease. Furthermore, it was necessary to determine whether the chosen method would allow for comparison among the various communicable diseases in different countries, and possibly also with other chronic diseases.

1 Decision No 2119/98/EC of the European Parliament and of the Council of 24 September 1998 setting up a network for the epidemiological surveillance and control of communicable diseases in the Community

2 http://ecdc.europa.eu/en/publications/Publications/1011_SUR_Annual_Epidemiological_Report_on_Communicable_Diseases_in_Europe.pdf

Metric selection based on literature workshops and expert meetings

In order to carry out these necessary steps and to obtain the most appropriate method, a literature review was performed, focusing on methodological and modelling issues related to both the measurement of burden of disease in general, and that of communicable disease in particular. The findings were discussed at workshops organised in close collaboration with ECDC. Participants included all BCoDE consortium members and other experts, with expertise ranging from epidemiology, health economics and demography to public health policy-making. The conclusions of the review and discussion process form the basis for the methodology chosen for calculating the first disease burden estimates under the BCoDE project. The methodology is set out in this protocol.

Further information necessary for disease burden calculations

In addition to choosing the methods for calculating disease burden from a variety of data sources, we collected information concerning data availability, disability weights and parameters describing disease progression. This information is necessary for the actual disease burden estimates. A literature review was conducted in order to define the potential health states for both acute illness and sequelae based on association and causal links. An inventory was compiled of the data sources available in the EU Member States and EEA/EFTA countries and a scoring exercise was carried out to analyse the strengths and weaknesses of these (e.g. notification data, hospitalisation data, surveillance data). Given that most of the data are far from complete, a literature review was conducted to define potential methods for correcting under-reporting, over-reporting and under-ascertainment.

Objective

The primary objective of Work Package 1 was the development of a methodology to generate evidence-based, valid and comparable estimates of the burden of communicable diseases and related conditions on society within EU Member States and EEA/EFTA countries.

Methodology for the BCoDE project

Selection of pathogens and countries

Countries considered

Disease burden calculations of the selected pathogens are planned for all European Union Member States and EEA/EFTA countries.

Selection of pathogens

Inclusion/exclusion criteria for the selection of pathogens were submitted to an ECDC ad hoc Advisory Forum working group and a final list of pathogens was agreed upon.

The aim of the BCoDE project is to estimate the burden of disease for as many as possible of the 49 diseases listed in Decision 2119/98/EC. However, a more formal and systematic procedure was deemed necessary in order to obtain a final, validated list of diseases. ECDC therefore developed an explicit procedure to identify the final list of pathogens/diseases to be included in the study.

A list of inclusion/exclusion criteria was submitted to an ad-hoc ECDC Advisory Forum working group for scoring. For methods and the final list of diseases included in the BCoDE project, see Annex 1.

Methodology choices for calculating disease burden

The methodology choices

The term "burden of disease" (BoD) is used to quantify the impact of a disease on a geographical region or population. However, since diseases affect patients and populations in various ways, a multitude of indicators can be used to quantify the BoD (e.g. incidence estimates, prevalence estimates, Disability-Adjusted Life Years [DALY]). Given the aim of the BCoDE project, the rationale for choosing a BoD indicator was that it should capture and weigh the distinct symptoms, severities, chronic sequelae and incidence of morbidity and mortality associated with each hazard in a single metric. A further rationale was that this integrated measure should allow comparison between communicable diseases within and among countries and, where possible, comparison with burden related to other conditions (not caused by communicable diseases). Finally, the chosen indicator would have to take into consideration specific issues related to communicable diseases and be sufficiently flexible to allow future development towards a methodology that can account for the dynamic nature of infectious disease epidemiology and the impact of intervention.

Integrated measures

In this project, a non-monetary integrated measure for health valuation will be used which can act as the basis for economic evaluations at a later stage.

Both monetary and non-monetary integrated measures have been developed to value health effects. These measures have in common that they integrate the health state associated with a particular disease in several countries into a single metric.

In the BCoDE project the focus is on the burden in terms of morbidity and mortality and not economic evaluation or assessment. For this reason non-monetary integrated measures will be used here which can, at a later stage, provide the basis for an economic analysis if desired.

DALY and QALY: the most prominent non-monetary integrated measures

This project will apply the DALY methodology to measure health gaps.

Among the various non-monetary integrated generic measures (i.e. measures that can be used across different diseases) in the scientific literature, Quality-Adjusted Life Years (QALY) and Disability-Adjusted Live Years (DALY) are the most prominent and they are also the most appropriate measures for disease burden calculations of communicable diseases.

The QALY approach assigns a quality-of-life index (i.e. a utility) to each health state, reflecting the desirability of that health state. Health states are usually valued between 0 (for death) and 1 (for perfect health). The QALY loss associated with an adverse health state is measured as the difference between QALYs with and without the condition. An overview of the different features specific to the QALY methodology is given in Table 1. QALYs are widely used in economic evaluation. The utility weights used in these studies are generated either directly or indirectly. If elicited directly, the preferences of either the patient, the general public or the physician are obtained using one of the following techniques: Standard Gamble (SG), Time Trade-Off (TTO), Person Trade-Off (PTO) or Visual Analogue Scale (VAS) (see Box 1). However, the choice of preferences – the values of patients, physicians or the general public – may influence the results. Some studies indicate that patients who have experienced the disease and partially adapted to it tend to place higher values on dysfunctional health states than the general population. This discrepancy becomes even more marked when patients value their own health state. The main factors contributing to the differences observed are poor descriptions of health states (for the general population), use of different internal standards, response shift and adaptation. In the case of indirect utility assessment, the dimensions of specific health states are rated on the basis of validated questionnaires (index instruments), such as EQ-5D, SF-6D, HUI and 15D. They are then validated using general population preferences. The algorithm for assigning values to each health state is obtained from a sample of the general adult population using one of the valuation techniques described in Box 1 (e.g. SG, TTO and VAS). However, these different instruments estimate different utility weights for the same health state, making it hard to compare them.

The DALY belongs to the family of health gap measures that calculate health losses based on the gap between the current health status and some ideal health goal that is defined arbitrarily. Thus, one DALY represents a loss of one year of life lived in perfect health. In the DALY methodology the severity of disease states is quantified by the use of disability weights that place a value on disease states between 0 (perfect health) and 1 (death). An overview of the different features describing the DALY methodology is given in Table 1. The DALY methodology was jointly developed by the World Bank, the World Health Organization (WHO) and the Harvard School of Public Health in the late 1980s for the Global Burden of Disease (GBD) study. The main idea behind the framework of the DALY was to incorporate both mortality and non-fatal health outcomes into a single measurement unit. This unit was to provide a comprehensive and comparable tool for describing the burden of disease and conditions in all countries worldwide. To meet this objective, the basic assumption was that similar events would be treated equally in all populations to ensure comparability. For example, the loss of a finger in Zimbabwe and the loss of a finger in Turkey should contribute equally to burden of disease. The original GBD DALY measure, its components and methodology were debated in the literature and in various international forums. The methodology has now been established as the gold standard and is applied in various national and sub-national burden of disease studies (see Table 1 for examples). The GBD project is ongoing and some major revisions have been made to the methodology. In the first GBD study, Murray and Lopez (1,2) used the PTO method (see below) to determine disability weights for more than 100 diseases and related sequelae. The weights were used in subsequent GBD studies in 2000, 2002 and 2004. For a future update of the GBD study WHO is currently revising the disability weights using novel elicitation methods. The main objective of the revision is to improve the methodology and to present a transparent, rigorous and standardised approach for generating disability weights. In addition to the GBD severity weights there are also some national severity weights available.

Table 1. Guidance checklist: An overview of DALYs and QALYs

Disability-Adjusted Life Year (DALY) as developed by WHO	Quality Adjusted Life Years (QALY)
Summary measure of population health type Health gap.	**Summary measure of population health type** Health expectancy
Unit of measurement • Years of life lost in a population due to premature death and disability, referring to an arbitrarily predefined health goal • Sum of Standard Expected Years of Life Lost (SEYLL) and Years Lost due to Disability (YLD)	**Unit of measurement** • Is a product of survival time and quality of time • QALYs are generally expressed as QALYs obtained by comparing two populations
Health dimension under consideration • Disability– i.e. loss of functional capacity	**Health dimension under consideration** • Health-related quality of life associated with certain health states
Information on mortality • Yes - SEYLL	**Information on mortality** • Yes - standardised life expectancy of the population under study (i.e. general population; drug-users, etc.)
Information on non-fatal health outcomes? • Yes - YLD	**Information on non-fatal health outcomes?** • Yes - through quality of life associated with non-fatal health outcomes
Disease-specific approach? • Yes - e.g. conditions can be linked to an ICD classification	**Disease-specific approach?** • No - approach based on health state descriptions which might be indirectly linked into a disease-specific approach
Characteristics of the measurement • DALYs developed for assessments in the Global Burden of Disease study • Internally consistent set of DALY estimates • Comparability of DALY estimates between populations and diseases and over time • Comprehensive set of disease and injury causes, in principle nothing is left out • Includes 108 specific disease and injury conditions, classified in a tree structure with four levels of disaggregation. First level comprises Group I, II, and III conditions: Communicable diseases, non-communicable diseases and injuries • Like events are treated equally – i.e. like events contribute to the same number of DALYs, irrespective of the individual's environment • Individual characteristics are restricted to age and sex	**Characteristics of the measurement** • Developed by economists, decision scientists and psychologists • Preference-based measure • Intended use for evaluation and intervention planning • Comparison of interventions • Primary use for cost-utility analysis • Applies cost-utility ratio which describes the incremental price of obtaining a unit of health effect • Utilities may differ from country to country/region to region • Utilities may be derived directly, or indirectly
Data requirements • Mortality data, life table, population data, causes of death • Incidence, disability weights (currently only available for the Netherlands, Estonia and WHO), duration of disease	**Data requirements** • Expected duration of stages; time-to-event or event probabilities and utility weights; (all factors of a natural history model)
Value choices • Age weighting • Sex-specific weights • Time discounting • Disability weights • Health goal	**Value choices** • "If... then" applied, also age weighting • "If... then" applied, sex-specific weights • Time discounting • Utility weights
Example used in: • WHO Global Burden of Disease studies • More than 20 national burden of disease studies (e.g. USA, Mexico, Chile, Turkey, Brazil, Australia, Singapore, Iran, Peru, Netherlands, South Africa, Zimbabwe) • Several local burden of disease analyses • Disease and risk factor-specific assessments (e.g. chikungunya, dengue, food-borne pathogens, lead, vitamin A, stroke, hepatitis C, climate change)	**Example used in:** • Routinely used in economic evaluation of medical care, technology and public health interventions in Europe and elsewhere
Limitations • No co-morbidity • Comparative risk assessment restricts the attribution of DALYs to single risk factors.	**Limitations** • Value sets derived from different countries are not comparable and cannot be used for cross-national comparisons • Resource and time intensive.

Both summary measures are suitable for quantifying the burden of communicable diseases and make it possible to compare the burdens of different communicable diseases within and between countries. Furthermore, DALYs and QALYs allow disease-specific weighting of health consequences. Although QALYs are commonly used in economic evaluations, the utility weights in these studies come from different sources employing various instruments and techniques. As economic evaluations are mostly conducted at the national level, the utility

weights usually represent national estimates, sometimes making a comparison between different QALY-based studies difficult. It would be difficult or even impossible to compile a list of utility weights based on the same underlying assumptions and using the same techniques for all health states related to the communicable diseases under study in the BCoDE project. The advantage of using DALYs is therefore that disability weights obtained using the same methods and assumptions are available from the GBD study for a large number of communicable diseases and their sequelae. Using them will to some extent ensure comparability with the GBD estimates. Therefore, at this stage of the BCoDE project, the DALY methodology is the most suitable for disease burden calculations.

Disability-Adjusted Life Years

The DALY is a health gap measure that extends the concept of potential years of life lost due to premature death to include equivalent years of "healthy" life lost in states of less than full health, or in more general terms disability. One DALY is one lost year of healthy life (WHO definition). The DALY methodology has been described by Murray et al (1-2) in the Global Burden of Disease (GBD) project using the following equation:

$$DALY = YLL + YLD.$$

YLL is the number of years of life lost due to premature death and YLD is the number of years of life lost due to disability. The number of years lost due to disability are computed by weighting each remaining life year with a factor between 0 (perfect health) and 1 (death) for an incident disability. The YLL due to a specific disease in a particular population is calculated by totalling the number of all fatal cases due to the health outcomes of a specific disease, each case multiplied by the remaining individual life expectancy at the age of death. More specifically, let $n_i(a)$ be the number of incident cases at age a of a disease with health states $i = 1,...,k$, where k is the number of different health outcomes due to the disease, and $d_i(a)$ the number of deaths at age a due to health outcome i. If $E(a)$ is the remaining life expectancy at age a, and a_{max} the maximum age, then:

$$YLL = \sum_{a=0}^{a_{max}} \sum_{i=1}^{k} d_i(a) E(a)$$

The YLD is calculated as the product of the duration of the illness and the disability weight of a specific disease outcome accumulated over the number of incident cases for all health outcomes. If the duration of health outcome i is indicated as D_i and its disability weight as w_i, then:

$$YLD = \sum_{a=0}^{a_{max}} \sum_{i=1}^{k} n_i(a) D_i w_i$$

$E(a)$ and D_i have time as a unit, all other quantities are dimensionless numbers. Age is discretised into (a_{max} +1) yearly age classes.

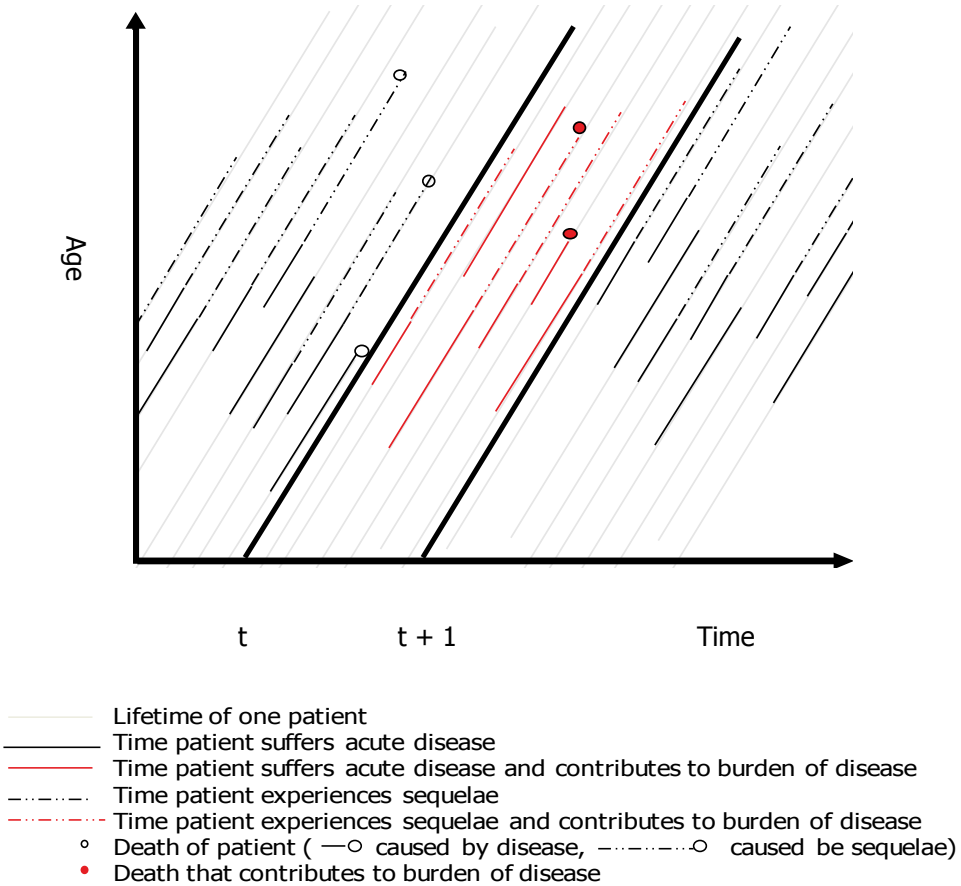
Pathogen-based DALY approach

Different methods have been developed for computing DALYs. DALY computations depend on the underlying assumptions as to how sequelae are linked to one another and to their infectious causes, how transitions between disease states are defined and quantified and whether time since infection is explicitly considered when describing disease progression. In the DALY approach applied in the GBD project (1-2) (3-4) disease burdens were estimated by totalling years of life lost (YLL) due to premature deaths and years of life lost due to disability (YLD) from incident disease cases. For some diseases long-term sequelae were included in the estimates, but many long-term sequelae were treated as separate disease entities with diverse aetiology.

For communicable diseases, the burden of disease associated with long-term sequelae is in some cases higher than the burden of disease associated with acute illness (e.g. STEC O157, Campylobacter). In order to attribute all the health consequences of an infection to the infection event, and therefore estimate the full burden caused by the infection, sequelae have to be considered as consequences of the infection. This is done by adopting an approach known as the pathogen-based DALY approach, as suggested by van Lier et al. (5). This approach, which relates all sequelae to their infectious causes by means of an outcome tree, will be used in the BCoDE study.

Relating sequelae to their infectious causes by means of an outcome tree involves the description of all health outcomes and their possible transitions in a flow chart. An individual enters the flow chart by infection and leaves by recovery or death. All intermediary states contribute to the disease burden. In contrast to full disease progression models that describe events after infection in relation to the time since infection, outcome trees do not take the time element into account. In a sense, an outcome tree attributes the complete disease burden to one point in time – the time at which infection occurs. Therefore, the outcome tree method is not suitable for projecting future burden estimates (Annex 2). With long-term sequelae projected to the present and treated as acute events, the crucial underlying assumptions of this method are that demographic and disease dynamics remain a steady state (see Figure 1). Although from a methodological point of view a full disease progression model is preferable, we realise that this cannot be achieved within the scope of this project. We have therefore chosen to base our estimates on outcome trees with the possibility of extending our estimates to full disease progression models in the future. The "pathogen-based DALY approach", as used in the pilot study by van Lier et al. (5) will be the first method implemented in the BCoDE project. Extensions of the method may be considered at a later stage.

Figure 1. Graphical illustration of the pathogen-based DALY approach



Incidence approach and pathogen-based approach

The basis for burden of disease estimates will be incidence estimates of infection by pathogen.

Pathogen-based approach

Disease burden estimates can either be assigned to an outcome – i.e. clinically defined categories of diseases irrespective of their cause, or to a cause – i.e. a pathogen. For communicable diseases we considered it important to assign the disease burden to a particular pathogen, as this makes it possible to assess the disease burden of both acute illness and related sequelae/complications caused by a particular pathogen. The pathogen-based approach will therefore be applied in the BCoDE project.

Incidence approach

Disease burden calculations can be based on incidence of infections or on incidence or prevalence of health outcomes. Premature death can also be viewed as a health outcome. The difference is that incidence of infection is measured per unit of time whereas the prevalence is a number or proportion. An incidence approach implies that all health outcomes, including their durations and weights, will be attributed to one point in time, when the infectious event takes place. From a practical point of view, there is a tendency to use the data notified in surveillance systems as a baseline incidence. Having decided to apply the pathogen-based approach, for which the full array of sequelae related to communicable diseases are assigned to the pathogen, the incidence approach is the most appropriate method and will therefore be applied in the BCoDE project. However, this does not preclude the use of prevalence data to estimate the number of chronic infections or other sequelae arising from an infection. For example, prevalence data will be used to validate results from incidence-based estimates if information on attributable fractions by cause is available for an outcome.

Outcome trees

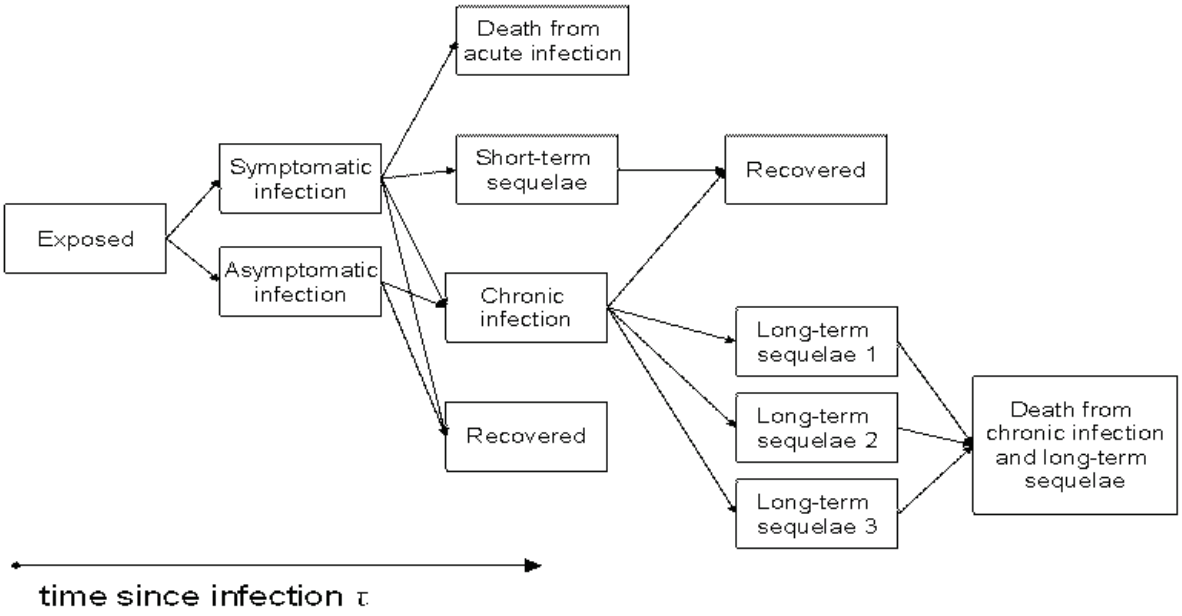
Sequelae will be linked to their causative pathogen by means of outcome trees. In this way burdens caused by sequelae will be linked to their causes. For validation purposes, attributable fractions will be used to interpret prevalence data for outcomes with competing causes so that overall burdens are not overestimated/counted twice. Outcome trees are based on the Bradford-Hills criteria for causality and on recommendations and validation from ECDC Disease-Specific Programmes.

Possible health outcomes of infectious diseases range from acute self-limited disease to chronic disabilities or even death. In order to assess the burden of disease for the selected pathogens, we need to define the different health outcomes following infection by a particular infectious agent. These disease outcomes can be described in the form of an outcome tree (see Figure 2). An outcome tree gives a qualitative representation of the disease progression over time by ordering all relevant health states following infection and illustrating their conditional dependency.

The construction of outcome trees implies making choices on the outcomes to include or exclude from the analysis. The first inclusion/exclusion criterion would be the strength of evidence for a causal relationship. According to Mitchell et al (6) examples are: 1) Sufficient evidence of a causal relationship 2) Sufficient evidence of an association 3) Limited or suggested evidence of an association 4) Inadequate or insufficient evidence to determine whether an association exists and 5) Limited or suggestive evidence of no association.

Outcomes that contribute little to the final result (because they are extremely rare and/or low-severity) could be excluded. Furthermore, the construction of the outcome tree is usually also guided in part by data availability. As scientific knowledge grows in the future, new outcomes may be causally linked to particular pathogens and the outcome tree may need to be updated, possibly increasing the burden of disease attributable to a particular pathogen.

Figure 2. An outcome tree describes all outcomes caused by an infection



An outcome tree was drafted for each of the selected pathogens, including all outcomes with sufficient evidence of a causal relationship or an association (first two categories). These outcome trees were then assessed by the Disease-Specific Programmes (DSPs) at ECDC and finalised, incorporating modifications based on the experts' recommendations.

Value choices

Applying the DALY methodology involves decisions which should reflect relevant value choices, as described below.

Life expectancy

Standard Expected Years of Life Lost (SEYLL) will be used as the baseline. Our approach will be modular to allow replacement of one life expectancy distribution with another.

Estimating YLLs requires the definition of a population health goal. To a certain degree the decision for a health goal is arbitrary and depends on the objectives of a burden of disease study. The health goal, defined in terms of healthy life expectancy, represents a point of reference. A burden estimate is therefore always a measure of a population's deviation from that reference point. A health goal may represent a potential limit to healthy lifespan or a given life-expectancy in ideal health. YLLs are then measured as the difference in years between the age-dependent health goal and the age of death.

As in the GBD study, we will use the most recent Standard Expected Years of Life Lost (SEYLL), a life expectancy distribution based on that of the Japanese population. In Japan, life expectancy at birth is the highest worldwide for both men and women. Alternatively we will consider a European life expectancy distribution, as suggested by ECDC's internal management. The Icelandic life expectancy distribution for males and the French life expectancy distribution for females are the highest observed in the European Union and EEA/EFTA countries (Eurostat)[3].

[3]http://epp.eurostat.ec.europa.eu/statistics_explained/index.php?title=File:Life_expectancy_at_birth(years).PNG&filetimestamp=20090430100004

However, since the mortality burden of communicable diseases is relatively low, using a common life expectancy for different European countries might overestimate the relative impact of communicable diseases in countries with a lower life expectancy due to chronic causes. For these countries, comparison with burden estimates based on national life expectancies can be useful to assess the magnitude of overestimation. In this project, we will follow a modular approach that allows the replacement of one life expectancy distribution by another. This will also enable scenario analysis with respect to health goals.

Disability weighting

WHO-generated disability weights will be used. For health states where weights are not available weights will be used from health conditions of similar severity. As updated WHO weights become available these will be used as the gold standard.

Disability weighting means that each outcome of a disease is assigned a value on a scale from 0 (perfect health) to 1 (death). These values may be dependent on the population in which they are measured, reflecting the norms and life conditions of that society. Ideally, the disability weights used to estimate a burden should reflect the values measured in the populations studied. However, disability weights are not usually available at the national level. Disability weights can be determined using different methods. Ideally, the weights reflect the values for the general population (as opposed to specific population groups) used for policy-making purposes at the national or international level. Weights based on elicitation panels made up of lay persons are increasingly becoming available, whereas in the past work was based on the results of expert panels (e.g. medical professionals). Values from patients who actually suffer from a disease are not considered informative for the purposes of estimating burdens, because patients' coping behaviour tends to influence the valuation process.

Different valuation techniques (see Box 1) are available for panel elicitation, including the Standard Gamble (SG), Time Trade-Off (TTO), Person Trade-Off (PTO) and Visual Analogue Scale (VAS). All methods give different results (usually VAS > TTO > PTO > SG), but they are correlated. The SG and VAS are sometimes problematic because they are only sensitive to severe disease (SG) or very sensitive to mild diseases (VAS) and this may lead to compression at either end of the scale. For this reason the TTO and PTO methods are often preferred.

The international transferability of disability weights is also a matter of concern. A study in western Europe concluded that there was "a reasonably high level of agreement on disability weights in western European countries with the VAS and TTO methods, but a lower level of agreement with the PTO method" (7). However, a recent study concluded that "Meaningful differences exist in directly elicited TTO valuations of EQ-5D health states between the United States and United Kingdom general populations" (8). Hence, severity weights should ideally be based on specific elicitations for the population under investigation, but this may be very difficult for the EU or even for a specific country. Given that the most comprehensive set of severity weights has been generated by WHO , and alternative weights are only available sporadically for specific countries (e.g. the Netherlands) or health states, the BCoDE project will use the internationally available disability weights from the GBD study. This enables us to compare our burden estimates with those obtained by WHO (assuming that we also use the same life expectancy distribution). WHO is currently updating its disability weights using a novel elicitation method to include health states that were previously not valued. By taking a modular approach to the computing of burden estimates we will be able to replace the present disability weights by new ones as they become available.

Box 1. The different valuation techniques

Standard Gamble (SG)

The SG method is the classic method for determining individuals' preferences, incorporating uncertainty as an element of the decision-making process (9). Using the SG technique, a scenario of a particular health state is presented to the respondents. They have to choose either to live with the presented health state for the rest of their lives or to receive some form of intervention, with an arbitrary risk probability of dying and another of getting well and living out their lives in optimal health. If the respondents decide to live with the presented health state, the gamble will be made more attractive by increasing the chance of getting well again. If they choose the gamble, the chance of getting well will be reduced. Having completed the first step, the respondents are repeatedly asked to select one of these two choices in different scenarios. After reaching a point of indifference, the respondents are unable to make further choices between the two possibilities. At this point it is possible to determine a utility scale number. This number determines the health scenario in the first alternative in relation to the health scenario in the second alternative. Thus, it is possible to determine sets of utility weights for different health states (10-12).

Rating or Visual Analogue Scales (RAS, VAS)

An alternative method to determine severity values for health states is to ask an individual to indicate the relative desirability of a state on a scale between 0 and 1, with 0 representing death and 1 a state equivalent to full health. Analogue rating scales are widely used to generate health state preferences. For example, the RAS and VAS technique has been used in the Quality of Well-Being Scale (QWB) and the EuroQol (EQ-5D) (10).

Time Trade-Off (TTO)

Another technique for assessing weights for different health states is the TTO method developed by Torrance et al (13). In this exercise respondents are asked whether they would like to spend the rest of their lives in a described health state, or to live a shorter life in a state of excellent health. As with the SG method, the respondents can choose between these two scenarios. If the respondents decide to live with the inferior health state for longer, the fraction of time will be prolonged in the alternative scenario. If the respondents opt to live a shorter life in excellent health the fraction of time will be reduced. As with the SG method, the continuous use of the TTO exercise leads to a point of indifference at which the remaining lifetime in the inferior health state becomes equivalent to less time in an excellent health state. This point of indifference is then used in an equation to calculate severity weights for different health states (10).

Person Trade-Off (PTO)

The preference values can also be determined using two variants of the PTO method. In the first exercise, as performed by Murray et al in the first Global Burden of Disease study (2-3), the respondents were asked to trade off the life extension of people at different health stages. They were first asked to trade off the life extension of healthy individuals and individuals in a given health state that was assumed to be less than excellent. In a second PTO exercise, the respondents were asked to decide between increasing the quality of life of people with inferior health or extending the life of people in excellent health for one year. The respondents were then asked to share their preference weights for the first and second PTO exercise in private. Afterwards all respondents were shown the results of the other participants' valuations and given the chance to discuss why they had chosen their preferences. After the discussion the respondents were able to revise their weights (2-3). A further form of the PTO method has been developed in the European context. This method involves finding a trade-off between preventing incident cases of a rapidly developing fatal disease and preventing incident cases of a chronic disease. Respondents are asked how many cases of chronic disease should be avoided before they can be indifferent to 100 rapidly developing fatal cases. If the number is 1000, the disability weight for this disease will be 0.1 (14).

As our outcome trees for the selected pathogens also include health outcomes for which WHO has not yet defined disability weights, we will have to "adapt" the disability weights from the existing list. This will be done by classifying both WHO existing health states and the new outcomes in accordance with the EQ-5D classification system. Each health state/outcome will be classified using five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression, and three existing levels per dimension. Altogether there will be 243 possible health states. Based on this "EQ-5D state" we will then match new outcomes with existing health states against the same "EQ-5D state". If there is more than one comparable health state we will determine which is the closest in terms of symptoms and preferences, based on the expert opinion of physicians.

Time discounting

Time discounting will not be applied in the base case estimates.

Discounting benefits and costs over time is a common concept in the field of economics that can be applied to both health and non-health effects. The principle of discounting is based on the common observation that people prefer benefits now rather than in the future.

The use of discount rates, or even differing discount rates, obviously has an impact on the burden estimates and may result in limited or no comparability among the generated DALYs and other studies. So far, the GBD estimates have been based on a three per cent time discount rate. However, several national and sub-national studies have not included any discount rate. Future GBD estimates will therefore be presented as both discounted and undiscounted estimates.

Time discounting cannot be applied to the DALYs used in this study, for which both acute illness and sequelae are linked to the pathogen and long-term sequelae are projected to the present. Discounting can only be applied properly when modelling the full time-dependence of infection and sequelae. Only when the occurrence of sequelae can be determined in time (since infection) can the time between infection and occurrence of sequelae be discounted. Therefore, for the approach adopted in the base case analysis discounting will not be used. This may change when disease progression models are available for some infections included in the study – a necessity when modelling the impact of interventions on disease burden. The implication is that in this study future and present disability are weighted equally.

Age weighting

Age weighting will not be applied in this study.

Integrating an age-weighting function into the DALY measure reflects the value given to life lived at different ages. This can be a subjective value but also an economically relevant value, reflecting the productivity of individuals of certain ages within society. Age weighting is highly debated and the exact quantitative implementation is controversial.

The decision to include age-weighting in the first GBD study was made in order to best fit the human capital theory. This ensured that a value was placed not only on the "intangible costs" of suffering, pain and premature death, but also on the indirect potential loss of income of the individuals in question. In the past GBD estimates were presented both with and without age weighting (e.g. WHO GBD 2001 data).[4]

In the BCoDE project, the burden estimates should only include the intangible costs of reduced health – i.e. premature death and disability. Age weighting will therefore not be considered in the base case analysis. If there are compelling reasons for including age weighting at a later stage of the project (possibly to link to other studies) the modular modelling approach will allow the rapid implementation of any such extension.

Data and data availability

Time span

To estimate incidence and mortality, data collected over a three-year time period (2005-2007) will be used. When estimating incidence of infection the point of departure will be notification data from routine surveillance systems. For each disease an assessment will be made as to whether incidence data from this time period is representative of the disease epidemiology, or whether there are underlying time trends that should be taken into account. In the event of longer time trends a decision will be taken as to whether to expand the time period to include 10 years of surveillance data. If good population-based studies are available for other time periods they will be considered as well.

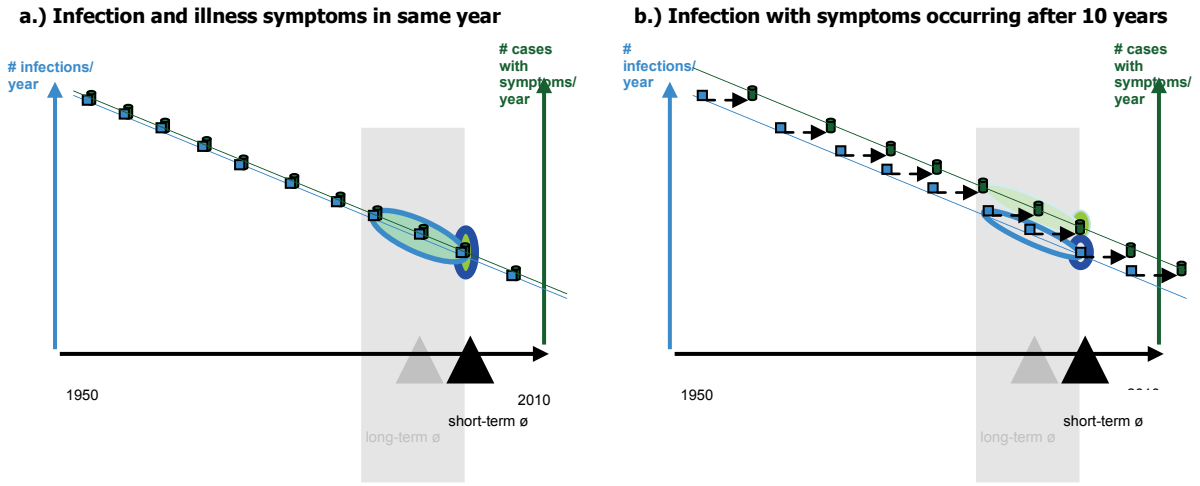
Communicable diseases differ in their long-term dynamic behaviour and may display distinct time trends in incidence over the years. While for some communicable diseases incidences remain constant for long time

[4] http://www.who.int/healthinfo/global_burden_disease/estimates_regional_2001/en/

periods, for others the incidence decreases (increases) or oscillates over time. If there is a monotone decreasing (or increasing) time trend in incidence, long-term averages would overestimate (or underestimate) the disease burden, as shown in Figure 3. Therefore, to avoid the influence of long-term trends on estimates we will base our incidence estimates on data from a three-year time period (2005–2007).

However, for infections with irregular (non-monotonic) time trends over one or several years, or for diseases that occur in incidental epidemics, a long-term average may be more representative of expected annual incidence. If a three-year time period is found to be too short to adequately represent typical incidences for a specific infectious disease, we will consider using a longer time period (e.g. a 10-year period). For infections taking the form of clearly defined outbreaks, the distribution of outbreak sizes will be considered in order to quantify expected numbers of cases and their variance.

Figure 3. Assuming a downward time trend for an infection displaying symptoms in the same year (a) and for an infection where symptoms only occur after 10 years (b)



- Blue rectangles represent the number of infections in the year of infection (exposure to an infection)
- Green "cans" represent the number of cases with symptoms where these cases occur in the same year as the infection (a) or a few years later (b), as indicated by the dashed arrow
- The long-term average (ø) is highlighted by a light-blue oval for incidence and a light-green oval for prevalence
- The short-term average (ø) is represented by a blue oval for incidence and a khaki/green oval for prevalence.

Number of fatal cases

The mortality data used will be stratified by age and sex, adjusted to correct for under- or over-reporting where applicable, and preferably obtained from national statistics institutes and EUROSTAT. The number of disease- and cause-specific fatalities from diseases with various causes is estimated either 1) by multiplying the absolute number of fatal cases obtained from disease-specific registers by an attributable fraction obtained from the literature, or 2) by modelling disease outcomes and mortality, taking into account the conditional dependency of the different health outcomes, as illustrated in the outcome tree.

Mortality data, preferably obtained from national statistics institutes and EUROSTAT, will be compared in order to estimate the burden due to premature death, stratified by age and sex.

One difficulty with this is that mortality is often not attributed to the underlying communicable disease that causes the condition leading to death. It is therefore necessary to estimate what fraction of the conditions registered as cause of death can be attributed to a communicable disease on the basis of cause-specific mortality. For some conditions and countries, estimates of attributable fractions are available. For others they have to be determined from outcome trees and information on disease progression and mortality rates.

The number of fatal cases from conditions with various causes, such as liver cancer (caused by hepatitis B and C virus infection or alcohol abuse) or Guillain-Barré syndrome (caused by *Campylobacter* or respiratory infection), are difficult to distribute by underlying cause. We need to know how the cause of death is recorded at death

registries in order to interpret mortality data and obtain an estimate of the number of deaths linked to a particular pathogen. In this project, the number of fatal cases caused by a specific communicable disease will be estimated by multiplying the absolute number of fatal cases from a condition, as obtained from disease-specific registers, with an attributable fraction for the disease obtained from the literature. Alternatively, we will estimate the disease-specific mortality by modelling the progression of disease through various outcomes starting from infection. This approach also takes into consideration the conditional dependency of the different health outcomes, as illustrated in the outcome tree. Ideally, both approaches should lead to comparable estimates for the number of fatal cases related to infection with a specific pathogen. If both estimates are available, one can validate the other.

Number of non-fatal health outcomes

Incidence data for number of infections, sequelae and mortality will be stratified by age and sex, adjusted to correct for under- or over-reporting where applicable, and preferably obtained from national health institutes or ECDC (TESSy). The number of cases of disease- and cause-specific sequelae are estimated either by multiplying the absolute number of cases, as obtained from disease-specific registers, with an attributable fraction obtained from the literature; or by modelling disease progression and occurrence of outcomes taking into account the conditional dependency of the different health outcomes, as illustrated by the outcome tree.

Depending on the communicable disease, incidences have to be estimated for a varying number of non-fatal outcomes. If incidence of infections and an outcome tree are available these estimates can be obtained by using the incidence of one outcome (e.g. number of symptomatic infections) at the root of the tree and the (conditional) probability of progressing to the next stage or to recovery. However, data to estimate the progression might not be available for complete outcome trees and supplementary data will be required. In addition to routine surveillance data collected for notifiable diseases the results of population-based epidemiological studies are required. Comparing notification data with evidence from population-based studies will provide insight into the quality of surveillance and possible levels of under-reporting and under-ascertainment. Incidence data should be differentiated according to surveillance level: non-consulting cases, cases where a PCP (primary care physician) was consulted and hospitalised cases (see Figure 4 below). This type of differentiation may help to ensure that the data is less biased by under-reporting, however it will seldom be complete. It is necessary to estimate the extent of under-reporting and of possible biases (from epidemiological studies) in the data. The degree of under-ascertainment may be available from community cohort studies or outbreak studies. The above implies that incidence of disease outcomes may be estimated from routine surveillance data, with correction for under-reporting and under-ascertainment, and by applying quantitative knowledge about disease progression. Alternatively, the incidence of disease outcomes has to be estimated from sources other than notification data (e.g. hospitalisation data together with information about attributable fractions), or information has to be taken from population-based studies. In a situation where the disease is in a steady state, or where infectious diseases develop within a short time span, these two approaches should lead to the same results. This makes it possible to validate estimates for some communicable diseases and their sequelae.

In the BCoDE project we will use surveillance data for notifiable infectious diseases (in particular TESSy, which is based on national surveillance systems), primary care data (where available), hospitalisation data and mortality data, all widely collected in European countries on a regular basis. Additionally, we will gather data from disease-specific databases such as cancer registries. Although collected regularly, these data sources are inconsistent and vary in quality depending on disease and country. Potential drawbacks are the definition used, the fact that nationwide coverage might not be given and the possibility that the collected data might only represent the more severe cases. Despite these drawbacks, the data sources are informative and do give indications of time trends and other related health problems. Nevertheless, for most of the communicable diseases under investigation these data sources are insufficient and only represent part of the iceberg. Corrections will be necessary for under-reporting. However, there is no single method applicable to all communicable diseases. The most suitable method for correcting under-reporting will have to be decided on the basis of each specific disease and country, depending on the evidence gathered and the data availability (of regularly collected and other sources). To this end, data other than notification data are needed and will be used where available.

Many non-fatal health outcomes can be caused by several pathogens or by non-communicable diseases, e.g. liver cancer and Guillain-Barré syndrome. Incidences related to a particular pathogen will be estimated by multiplying the number of reported cases with an attributable fraction obtained from the literature and/or by modelling the steps through the outcome tree, starting from incidence of infection and taking into consideration the conditional dependency of different health outcomes.

Duration and severity of health states

The duration of health states is based on national estimates, where available. This will be compared with estimates extracted from the literature. If no national figures exist, estimates from the literature will be used.

Durations of health states for chronic diseases are usually measured in years. Disability weights are then defined per life year lived with this disability. For infectious diseases with a short acute phase, this approach may not be appropriate. For short duration diseases disability weights can be determined per episode by focusing only on the phase of acute disease (period profile) or by focusing on a year in which an episode of acute illness is experienced (annual profile). Both methods have been tested and it appears that using the annual profile may overvalue disability weights. In the BCoDE project disability weights per episode (period profile) will be used where available and applicable.

For conditions with less than one year's duration the annual profile will be used for weighting. For conditions lasting more than one year the disability weights will be multiplied by the duration in years to obtain the total burden. Depending on the infection, health care systems also have an impact on the (conditional) probability of progressing to the next stage of care. Therefore, the duration of health states will be based on national estimates, if available. If national data is not available, estimates will be obtained from the literature.

Under-reporting

Correction for under- or over-reporting and under-ascertainment has to be evaluated specifically for each disease and country. Adjustments depend mainly on data availability and on disease-specific characteristics of diagnosis and reporting.

Introduction

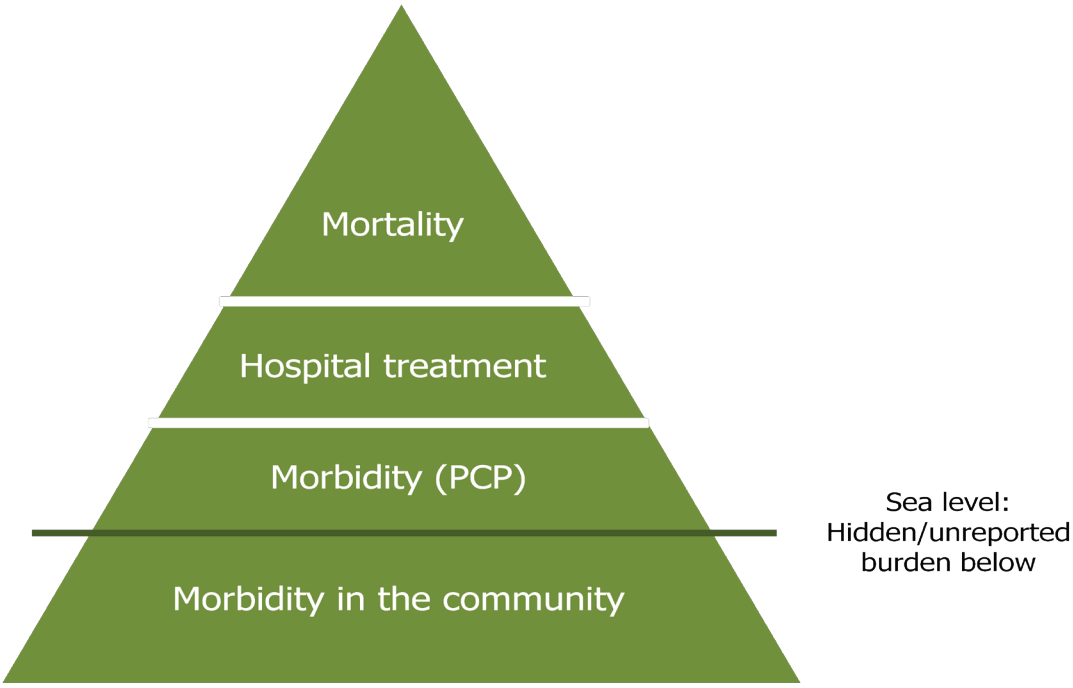
When estimating disease burden the under-reporting and under-ascertainment of data generate a degree of uncertainty. During a recent pilot study of the burden of infectious diseases in Europe, Van Lier et al suggested that the true burden of disease was not fully represented by raw, non-manipulated datasets owing to the probability of under-estimation for all diseases and countries (5). It would therefore appear to be crucial to identify areas of under-estimation in order to employ correction methods and determine the true burden of disease.

In the past there has been some confusion over the meaning of the terms under-ascertainment and under-reporting. In this study under-ascertainment is used to refer specifically to cases or exposure in the community which are not recorded by a notification or surveillance system due to the fact that healthcare advice is not sought. Under-reporting refers more specifically to cases where healthcare advice is sought but the infection status is misdiagnosed, misclassified, miscounted or the information summarised, meaning that full details are not passed on to national statistical bodies. In short, under-ascertainment occurs within the community and under-reporting occurs within institutes and involves physicians, hospitals, laboratories, governments and networks.

The 'iceberg' analogy

The surveillance pyramid or 'iceberg' metaphor (see Figure 4) illustrates the impact of infectious diseases on society and suggests patterns of health-seeking behaviour. More importantly, it indicates general levels of surveillance and data collection within the healthcare system, areas that are subject to varying rates of under-reporting and under-ascertainment. These levels include the general population (defined as the total population, including hospitalised and otherwise institutionalised persons in which all exposure and cases occur); PCPs and hospitals (which represent a proportion of all exposure) and the mortality level (representing a proportion of those hospitalised and a smaller proportion of those exposed). As under-reporting and under-ascertainment rates vary at each level, a correction must be applied separately for each level. Therefore, it is important to separate these different levels in burden of disease studies.

Figure 4. Burden of infectious disease in society — the 'iceberg'



Source: PPT adapted from a figure produced by the Health Protection Agency, UK (15).

Available methods - general

There are different ways in which to adjust data for under-reporting and under-ascertainment. All methods aim to estimate multiplication factors that relate reported data to complete and partially unreported data. In general, multiplication factors are age-dependent. The method used to estimate them depends on the specific disease, the type of data available and the reasons for under-reporting and under-ascertainment.

Multiplication factors

Multiplication factors can be applied to the reported number of cases of a particular disease in order to estimate the true number of cases.

Multiplication factors are developed by comparing incidence or exposure in the general population (determined by community-based or serological studies) with the data on notified cases (including incidence of hospitalisation, laboratory-confirmed cases, PCP-confirmed cases and deaths attributable to the disease). Multiplication factors must be disease-specific (since under-reporting exists in varying degrees for different diseases); country-specific (owing to variations in disease exposure, healthcare systems, availability of treatment or cultural, social and technological differences); age-specific (since rates of under-reporting differ among age groups for many diseases) or possibly gender-specific. In some instances multiplication factors may even need to be strain-specific. One example might be seasonal influenza where diverse seasonal strains will cause varying degrees of symptomatic disease and associated health-seeking behaviour. Thus, the community burden for strains causing milder symptoms and lower rates of healthcare-seeking behaviour will be under-represented to a greater extent. There are few community-based studies enabling the estimation of multiplication factors in such detail. In most cases we only have rough estimates of the ratio of reported to unreported cases.

In a study by Mead, multiplication factors were used to correct for under-reporting or under-ascertainment at the community level in hospitalisation data and mortality data (16-17). Multiplication factors borrowed from previous studies were used to estimate true illness in the community, whereby the degree of under-ascertainment of illnesses caused by *Salmonella* and other non-bloody diarrheal pathogens was 38-fold and the degree of under-reporting of illnesses caused by E. coli O157:H7 and other bloody diarrheal pathogens was 20-fold. For pathogens causing severe conditions the degree of under-reporting was 2-fold, since it was considered more likely that such cases would present to healthcare facilities. Hospitalisation data and mortality data were doubled to account for under-reporting, producing a lower figure than that applied to under-ascertained community data since it is believed that once in the healthcare system, there is a higher probability that cases will be reported.

In a second study by Van Damme et al. (18) a multiplication factor of six was applied to reported hepatitis B case data for all 30 EU and EEA/EFTA countries to calculate the 'true' incidence of hepatitis B exposure in the community. This was calculated by assuming a factor of two to correct for under-reporting and a factor of three to correct for asymptomatic cases, based on the knowledge that two-thirds of infections are asymptomatic.

Using the same multiplication factors for different countries and populations does not take into consideration vital inter-country differences, such as whether a disease is notifiable, whether the reported data has already been adjusted, how good the data collection and reporting systems are and whether there are differences in diagnostic approach between countries. In conclusion, when applying multiplication factors, the differences between countries have to be assessed and taken into account so that multiplication factors can be based on country-specific evidence. Obviously, multiplication factors are disease-specific and depend strongly on the clinical presentation and severity of an infection and how it is distributed in the population. Furthermore, clinical symptoms are usually age-dependent and may often be different for men and women. Rational, evidence-based choices have to be made for the values of all multiplication factors.

As a general rule, multiplication factors have to be separated into one factor describing under-ascertainment and one describing under-reporting. Both will be assumed to be age- and sex-dependent.

Ways of estimating multiplication factors

Statistical or dynamic models are used in almost all methods for estimating the extent of under-reporting and under-ascertainment. Statistical models are required to translate the results of telephone, postal or serological surveys into estimates of multiplication factors. In some studies, a modelling framework, rather than a statistical analysis, is used to translate clinical knowledge about the course of an infection and disease into a multiplication factor (e.g. a model relating the number of symptomatic cases to the total number of infections). Modelling frameworks may make use of reported prevalence, incidence or mortality data to calculate the expected incidence of specific outcomes. These can then be compared to the reported incidences to understand the magnitude of under-reporting and under-ascertainment.

By way of example, the decision tree model approach has been used to estimate under-detected cases of rabies in Tanzania (19) and *Trypanosoma brucei rhodesiense* sleeping sickness cases in Uganda (20). It is therefore assumed that with some modification of the model parameters the quantitative framework could be applied to a range of other diseases. For the project in Uganda, a deterministic model was also developed to estimate the proportion of undetected cases that sought healthcare.

Age-specific disease curves identify incidence of disease at a given age but can also identify ages for which there is significant under-reporting. Two studies that used this method also used regression as a secondary step (21-22). Another example is a study by Majowicz et al (23), where the authors used Monte Carlo simulation to estimate the number of non-typhoidal *Salmonella* gastroenteritis cases in order to calculate the global burden of this disease.

Some aspects of modelling (statistical analysis) will be clearly be vital for correcting under-estimation of disease burden. As this project progresses, it will be necessary to fully investigate existing and available models in order to decide whether they are suitable for the project, or relevant to the situation in Europe. It will also be necessary to discuss whether one disease-specific model would be appropriate for all European countries or whether country- or region-specific models would be more appropriate.

Models already exist for predicting or estimating the community burden of various diseases in a number of different countries. The use of existing disease models could be advantageous in that it would avoid the need to develop new models — a process which is often complex and time-consuming. However, it may prove necessary to develop new models in order to give the best estimate of disease burden and correct for under-reporting and under-ascertainment.

Capture-recapture methods and log linear models

The capture-recapture method involves comparing data from various sources and registries and establishing whether cases have been captured from more than one data source. To do this it must be possible to link the data at the individual level. This technique has been used in combination with analytical techniques, including log-linear models and others, to correct for under-reporting. Capture-recapture studies have been performed in the context of infectious disease surveillance for tuberculosis, meningococcal diseases and others.

Under-reporting of mortality can also be assessed by capture-recapture analysis. In a study of pertussis deaths in England (1994-1999), it was suggested that 46 deaths were attributable to pertussis but only 33 had been identified, classified correctly and reported to the relevant authorities (24).

Community-based studies

The umbrella term "community-based studies" includes serological surveys and population-based telephone, internet or postal surveys. Community-based studies aim to measure the incidence of infection (symptomatic and/or asymptomatic) in the population and in this way quantify the proportion of infected persons that contact the healthcare system and are diagnosed. They therefore provide a basis for estimating the multiplication factors used for under-ascertainment.

Population-based surveys

Population-based surveys actively search for community disease incidence, usually within a randomly chosen sample of the population (cross-section) or a cohort followed over a given time period. Typically this involves identifying cases that fulfil a case definition (gastroenteritis, influenza-like illness) and performing studies, including microbial diagnostics and embedded case-control studies, to investigate the healthcare-seeking behaviour that such cases provoke.

The community-based incidence described in such a study can be compared to that reported in notification systems, which should then give an idea of the magnitude of under-reporting and under-ascertainment. Population-based surveys are a useful way of gaining information on community-level under-reporting for those diseases that are self-limited or mild (e.g. gastrointestinal infections), and those that are asymptomatic and therefore do not prompt healthcare-seeking behaviour (some sexually transmitted infections (STIs), including gonorrhoea and Chlamydia). Population-based surveys can also be useful for estimating the under-reporting of infections diagnosed by physicians and in hospitals and laboratories. However, few population-based surveys are available because they are time-consuming and expensive.

Examples of population-based surveys

One type of population-based study is the cross-sectional telephone survey. This involves interviewing a randomly chosen sample of the total population to determine community opinions, habits, preferences and, in this instance, disease incidence. When investigating the occurrence of infections, syndrome information is usually requested. This involves determining whether an individual has experienced certain symptoms within a given time period, the severity of the symptoms, whether respondents sought healthcare and other information. In a recent retrospective cross-sectional telephone survey in Canada by Henson et al (25) the incidence of gastrointestinal disease was estimated. Other parameters, including hospital and physician visits, house calls, laboratory tests, medication and days of work missed, were useful for investigating the epidemiology of those diseases. The survey findings were first subjected to regression analysis and then modelling as a secondary step to estimate the monetary cost of burden.

A similar study was conducted by Majowicz et al (26) in a single city in Canada and by Thomas et al (27) in three locations across Canada, both trying to estimate the burden of gastrointestinal disease. Other examples of cross-sectional telephone surveys include a study that assessed the level of influenza vaccination coverage during two consecutive influenza seasons (2002/2003 and 2003/2004) in five European countries (28); a study which used computer-assisted telephone interviews to assess the lifetime and recent history of STIs and blood-borne viruses (BBV) in Australia (29) and studies of gastrointestinal diseases in USA (30) and Ireland/Northern Ireland (31).

The postal questionnaire is similar to the telephone survey with correspondence being conducted via post. Again the questions are oriented towards a particular outcome whereby certain answers would class an individual as a case (perhaps even a possible, probable or definite case) or a non-case. An example of a prospective community-cohort study combined with a PCP-cohort study was used in the United Kingdom (32-33) and in the Netherlands (34-35) to estimate gastroenteritis incidence at the different layers of the surveillance pyramid. Initially a baseline survey of an invited (randomly chosen) cohort was carried out and then weekly cards sent to report on any gastrointestinal symptoms. If the subject was suffering from diarrhoea, stool samples were also obtained so that pathogens could be identified. Conducting prospective studies, targeting specific syndromes and following up with etiological studies and registration of outcomes are among the most viable methods for obtaining information on incidence numbers, etiology, multipliers and outcomes, although probably the most expensive.

Another approach is the detailed analysis of large outbreaks and the use or non-use of medical care services in order to estimate the relationship between the levels in a pathogen-specific gastroenteritis reporting pyramid, for example. Combining such information with surveillance data gives an extrapolation of the number of cases in the general population. However, outbreak situations might not be comparable to sporadic cases in the population. Differences in susceptibilities of sub-populations, size of doses or other key areas can lead to differences in clinical symptoms and therefore in the likelihood of seeking medical care. Increased media coverage is a key factor, introducing biases in consultation behaviour during outbreak situations.

Face-to-face interviews, using a similar type of questionnaire to that described above, have also been used. Examples include a population-based study of Chlamydia in China (36) that used a computer-aided questionnaire and required participants to provide a urine sample, and a personal door-to-door survey of HIV and STIs in Slovenia (37).

Another type of population-based study that can help to correct for under-reporting is described by Hsieh et al (38). A cohort of patients diagnosed as having urethritis, cervicitis or gonorrhoea in various health centres in Taiwan were interviewed and retested for gonorrhoea which demonstrated that 42% of positive gonorrhoea cases had not been reported to the National Gonorrhoea Notification Disease System (NGNDS). The main reasons given by clinicians for not reporting were that clinical samples had not been obtained, the clinicians were afraid to violate patient privacy and that they were unaware the disease was notifiable.

Case-control studies in the population may also be used to estimate disease outcomes for specific pathogens. Case-control studies are performed for various purposes, such as determining risk factors or attributing illnesses to foods. They are usually conducted for laboratory-confirmed cases but can be applied to many of the approaches discussed above. Case-control studies are either prospective or retrospective and mostly rely on self-reporting. Self-reported symptoms should ideally be confirmed by medical examinations to correct for over-reporting.

Population-based studies have the advantage of finding unreported or unascertained cases, but there are also some limitations, the most important being the possibility of bias, of which there are many types. Sampling bias due to non-random sampling of a population could be an issue and this would result in the study not being representative of the entire population. Certain ethnic, migrant, age, occupational or other groups may not be included in the study because they are unregistered, hard to reach, do not have access to a telephone (in the case of telephone surveys), or cannot read and write in the language of the study. There also may be a response bias since only certain types of people will agree to participate in the study, which again may lead to the sample not being representative of the entire population. Measurement bias may also be significant in terms of the case definition. If the case definition is too general or specific this could greatly increase or decrease the incidence of disease. In addition, interviewer interpretation of the interviewees' answers may be biased or the respondents' memory of disease occurrence may be distorted. Finally, such studies are time-consuming and costly and are therefore not feasible for many countries and diseases.

Serological surveys

Serological surveys are a tool for investigating the immune status of a population. In a population-based serological survey serum samples are collected from a representative sample of the population. These can then be tested for the presence of antibodies against a variety of infectious diseases. Depending on the disease and antibody tested, the samples may give information about past and recent exposure and possibly vaccine-induced immunity. Usually this information can then be stratified by age and sex. In the case of diseases for which serological data gives information about recent incidence of seroconversions, comparison with notification and hospitalisation data reveals what proportion of seroconversions were diagnosed and reported. Under-reporting and under-ascertainment can then be estimated by comparing the number of sero-positive cases and hence exposed individuals (sero-prevalence) with the number of reported cases (incidence/prevalence).

Serological surveys have been conducted for a variety of infectious diseases, but only some of them are representative at the national level. Serological studies are a common epidemiological study method for estimating the prevalence of HIV and hepatitis B and C. For these diseases, acute infections are not easily detected and often remain asymptomatic. Similarly, they have a long chronic stage during which specific antibodies are present. Therefore measuring the population prevalence for these diseases gives the most reliable information about their epidemiology. Furthermore, serology can also act as a diagnostic tool for these infections.

Numerous studies have also been carried out for other diseases including viral haemorrhagic fevers (VHF) and Q fever, for which the gold standard of diagnosis is also serology. However, as these diseases are considered rare in most countries, serological surveys of VHF and Q fever tend to be conducted by research groups as opposed to public health institutes and organisations. Moreover, as a result of limited resources, study sample sizes are often small and limited to certain geographical regions.

Serological surveys have also been performed for diseases for which serology is a novel research approach. These include *Salmonella*, *Campylobacter*, *Brucella*, *Cryptosporidium* and syphilis.

Serological surveys among animal populations as a measure of animal disease and as a proxy for human disease have been carried out for *Salmonella*, *Brucella*, *Campylobacter*, tuberculosis, influenza, Q fever and others. Studies are mainly conducted by research groups but some data are collected by surveillance networks and interest groups including the European Surveillance Network for Influenza in Pigs (ESNIP) 2, and Med-Vet-Net (*Campylobacter, Salmonella*). The Med-Vet-Net studies were in humans.

Finally, serological surveys of vaccine-preventable diseases are common and ongoing. The aim is mainly to assess the immune status of the population, including vaccine-induced immunity, and to identify regions of low herd immunity and potential outbreak threat. It is often difficult to differentiate between antibodies generated as a result of natural exposure and those resulting from vaccination. Therefore the epidemiological interpretation of serological data in terms of incidence of infection is sometimes difficult.

An exception is pertussis, for which the impact of vaccination on antibody titres against pertussis toxin does not last long. Using information on the rate of decay of those antibodies after natural infection, the incidence of pertussis infections can be estimated from serological data. This was first done by de Melker, Versteegh et al. (39), resulting in an estimate for the proportion of pertussis infections notified per age group in the Netherlands. In a follow-up study, these methods were applied to age-stratified seroprevalence data from five European countries collected by the European Sero-Epidemiology Network (ESEN). The resulting incidence estimates show that there is a huge difference between the age distribution of notified cases and the age distribution of infections resulting in seroconversion.

If good information is available on vaccination coverage over time in a population, the fraction of the population that has experienced natural infection can be estimated from the serological profile. Moreover, a comparison with notification data will give some information about possible under-reporting and age-biases in reporting.

There are limitations to using serological surveys to determine under-reporting. Firstly, not all diseases have robust serological tests with reasonable sensitivity and specificity. Examples include pathogens with antibodies that cross-react with other pathogens such as *Listeria*, enterohemorragic E. coli and other bacteria. Thus, past serological surveys have been limited to a handful of diseases. For some diseases, extensive exposure may be needed to obtain a measurable serological response. For other diseases asymptomatic infections also result in sero-conversion. Interpretation of serological titres must therefore be disease and country-specific.

A second limitation is that antibodies originating from natural exposure to a pathogen cannot always be distinguished from antibodies resulting from vaccination. Therefore, serological surveys of diseases for which there is routine vaccination within the community, including tuberculosis (BCG vaccine), measles, rubella, and other childhood vaccine-preventable diseases, may not easily be interpreted. Furthermore, for some diseases it is not possible to distinguish between past and recent infections, meaning that for these diseases sero-prevalence represents a cumulative lifetime incidence.

Few sero-surveys have been carried out in representative population samples at the national level. Most serological surveys are limited to small samples of a population or to specific risk groups. Many serological studies have focused on measuring prevalence in high-risk groups (e.g. a serological survey of hepatitis C among intravenous drug-users). However, these groups would usually have a disproportionally high disease burden compared with the general population. Consequently, this data cannot be used to estimate disease burden for the population at large and supplementary information is needed concerning the size of risk groups and their contribution to overall transmission dynamics. In addition, different studies may use a variety of serological testing methods or different cut-offs to define positivity. This also limits comparability among studies.

Routine serological testing takes place in various settings, for example to ensure the safety of blood products from blood donations. Nevertheless, this information, which could be useful for assessing the population prevalence of diseases such as HIV, hepatitis B and C, is not systematically analysed and published.

There are European networks collecting and analysing serological surveillance data, including the European Commission-funded European Sero-Epidemiology Network 2 (ESEN2) which collects serological data on vaccine-preventable diseases and the European Network for Diagnostics of "Imported" Viral Diseases (ENIVD), a network collecting serological surveillance data on rare, imported viral infections including viral haemorrhagic fever (VHF). In the ESEN network, methods have been developed to standardise serological data enabling the comparison of serological profiles in different populations. The results of these seroepidemiological studies have been published for a large number of vaccine-preventable diseases.

Available methods and their potential application in the BCoDE project

General

The BCoDE project relies heavily on notification data and other readily available data from surveillance systems. Data from routine surveillance sources needs to be adjusted to correct for under-reporting and under-ascertainment. We distinguish at least three main types of data that represent the different levels of the surveillance pyramid and are obtained from different sources. These are:

- Incidence of symptomatic infections (morbidity data): Incidence of illness in the community which includes both those that sought healthcare assistance through a PCP and those that did not.
- Incidence of hospitalised cases of infection (morbidity data): Hospitalisation data should quantify the number of cases hospitalised and include additional details such as length of stay. Those hospitalised tend to be the more severe incidence cases.
- Incidence of death due to the infection and its sequelae (mortality data): The number of deaths is usually quite well reported, but there is considerable misclassification of the causes of death. There is under-ascertainment of secondary and tertiary causes of death. Moreover, details relating to conditions and possible sequelae that an individual died with (e.g. underlying causes) but not necessarily of (i.e. the primary cause of death) are frequently missing or not included in summaries.

By identifying and correcting for under-reporting and under-ascertainment at each level within the pyramid, a more accurate and valid basis will be obtained for estimating disease burden. The BCoDE project will, in the first instance, use data that is publicly accessible or available via national public health institutes, adjusting for under-reporting and under-ascertainment. The chosen correction methods will be disease-specific and the estimated multiplication factors will be country- and disease-specific, since the magnitude of under-reporting and under-ascertainment within the surveillance pyramid varies greatly among diseases and countries, or even within one country for different periods or regions.

Using current notification data and correcting for under-reporting and under-ascertainment will create a cross-sectional view of morbidity and mortality incidence and burden in the society. Cross-sectional data gives a snapshot picture of acute illness, sequelae and death at the present time. In contrast, the longitudinal view considers disease and its outcomes over time and may therefore provide important information on (long-term) sequelae of infection.

To obtain information about possible under-reporting or under-ascertainment of sequelae it will be necessary to utilise disease outcome trees. The steps required can only be performed after correction for under-reporting and under-ascertainment in the cross-sectional data. Briefly, the steps include:

- Adjust notification data for under-reporting and under-ascertainment.
- Enter the adjusted incidence data into the outcome tree model. Based on the transition probabilities given by the outcome tree, it is possible to estimate the proportion of incident cases that develop specific sequelae, recover or die from a particular complication. The model calculates the expected incidences of particular sequelae. Results are given in terms of incidence of sequelae and deaths.
- Expected incidences of sequelae can then be compared to reported data on sequelae and sequelae-attributable deaths. Differences between reported and expected outcomes will indicate under-reporting of sequelae, misclassification of sequelae or mortality and temporal changes in the epidemiology of the relevant disease.

Stepwise application process

A stepwise process for deciding how to deal with under-reporting and under-ascertainment and which correction method to use is described below. Ideally this process should be followed for each disease in each country to give country- and disease-specific correction methods.

Cross-sectional comparison (see Figure 5 below)

1. Gather information and evidence from national representatives.
   - *Contact national health authorities*:
     Their experiences could be useful for finding the most appropriate approach to correcting under-reporting and under-ascertainment. Furthermore, they may highlight the importance of the initial search for evidence if inter-country reported incidence data proves to be incomparable. It may be that a national

institute in one country publishes data already corrected for under-estimation, while an institute in another country only publishes non-manipulated data. It is vital to obtain this information from national health board representatives to be clear as to whether datasets are raw or manipulated.

- Ask the national health authority representative if the country applies any form of correction method to manipulate raw data before publishing or reporting to international institutes to correct for under-reporting and/or under-ascertainment.
- If yes, more detailed information must be sought: methods used to correct; reasoning behind the choice of methods; to which diseases are the methods applied and at which institutional level is the correction factor applied to data? Institutional level 1 refers to the local collecting system (PCPs, hospitals and other health units), level 2 refers to the regional collecting department; level 3 refers to the national collecting and reporting department/office and level 4 refers to the international collecting and reporting department, which may include the World Health Organization (WHO) or ECDC's TESSy. It is possible that there is some correction at each level and the extent of correction should therefore be determined for each level.

- *Contact National Surveillance Contact Points:*
  As during data investigation the National Surveillance Contact Point can be contacted. Similar questions may be asked and any information obtained can be recorded and sent to the central group.

2. Gather information and evidence from published literature to find past studies that correct for under-reporting or under-ascertainment or suggest true incidence/burden from which multiplication factors can be derived to correct under-reporting and under-ascertainment.

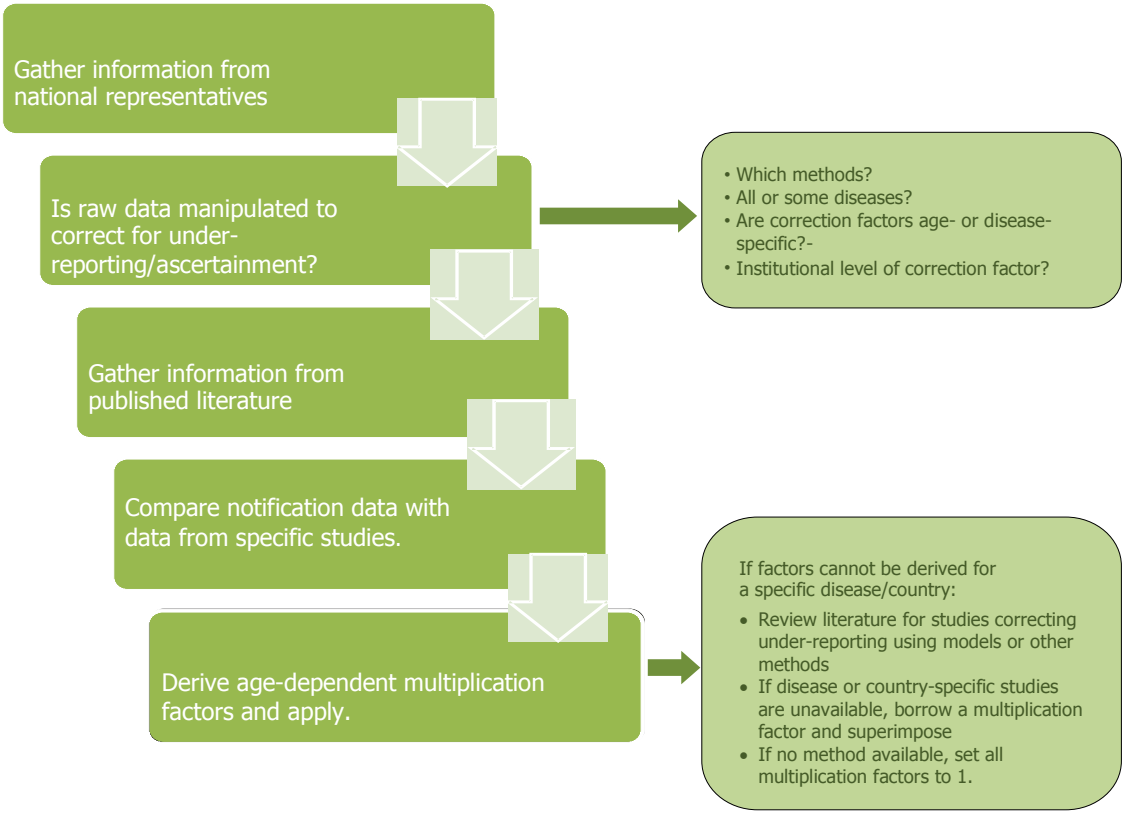
   An extensive, systematic review of the literature should be conducted to find studies that have been done at both the regional and the international level for each country and each disease under investigation. Comparisons between the data sources are essential in order to ascertain the level of under-reporting/under-ascertainment and establish correction methods.

   Study types likely to be most significant are:

- Studies identifying areas of under-reporting/under-ascertainment of data for the relevant diseases and countries
- Studies correcting for under-reporting/under-ascertainment for the relevant diseases and countries
- Community-based studies (e.g. serological surveys or community-based telephone surveys) provide information on the background level of exposure or infection and could indicate the magnitude of under-reporting/under-ascertainment. Multiplication factors can also be determined from this information in order to re-calculate reported data incidences for other datasets.
- Outbreak investigation studies: These studies are useful as an additional measure of incidence since they involve active surveillance. However, outbreak studies must be treated with caution since during an outbreak physicians and hospitals feel the need to report more and have more samples laboratory tested. Similarly, the public become more aware through media coverage and people are more likely to seek healthcare.
- Case-control and cohort studies: Useful for gathering information on development of sequelae following infection. Such studies may give more precise information about the rate and timing of sequelae development.
- Modelling: Studies that use modelling frameworks may take into account the dynamic aspects of transmission, the indirect effects of intervention and other outcomes resulting from the communicable nature of infectious diseases. Models can also be used to reconcile conflicting evidence or assess the strength of different data sources systematically. However, designing consistent, evidence-based models is time-consuming.
- Capture-recapture: This methodology can be used to estimate the completeness of registries and surveillance data. In particular, a capture-recapture study provides information about the proportion of cases not registered in any of the analysed data sources. In order to conduct a capture-recapture study there must be at least three largely independent data sources available for which data can be linked at the individual record level.
- Multiplication factors: Publications that have previously estimated multiplication factors for specific diseases to calculate corrected incidences from reported data will be useful as a guide to the magnitude of such factors.

3. Compare incidences, as estimated from notification data and other registries, with information from specific studies found in step 2 for cases of disease in the community, hospitalised cases and deaths.
4. Derive (age-dependent) multiplication factors for the three levels of reporting (notification, hospitalisation and mortality) and apply the multiplication factors to correct for under-reporting/under-ascertainment to obtain adjusted incidences.
   - If multiplication factors cannot be determined for the specific disease or country:
   - Review published literature to find past studies that correct for under-reporting/under-ascertainment using models or frameworks and assess their applicability to the current study.
   - Review published literature to find past studies that correct for under-reporting/under-ascertainment using any other method and assess applicability to the current study.
   - If a disease-specific or country-specific method cannot be found, a multiplication factor can be taken from elsewhere (a similar disease group or healthcare system) and superimposed for use in this situation.
   - If no applicable method of information is available, set all multiplication factors to 1.

Figure 5. Flowchart for cross-sectional under-reporting

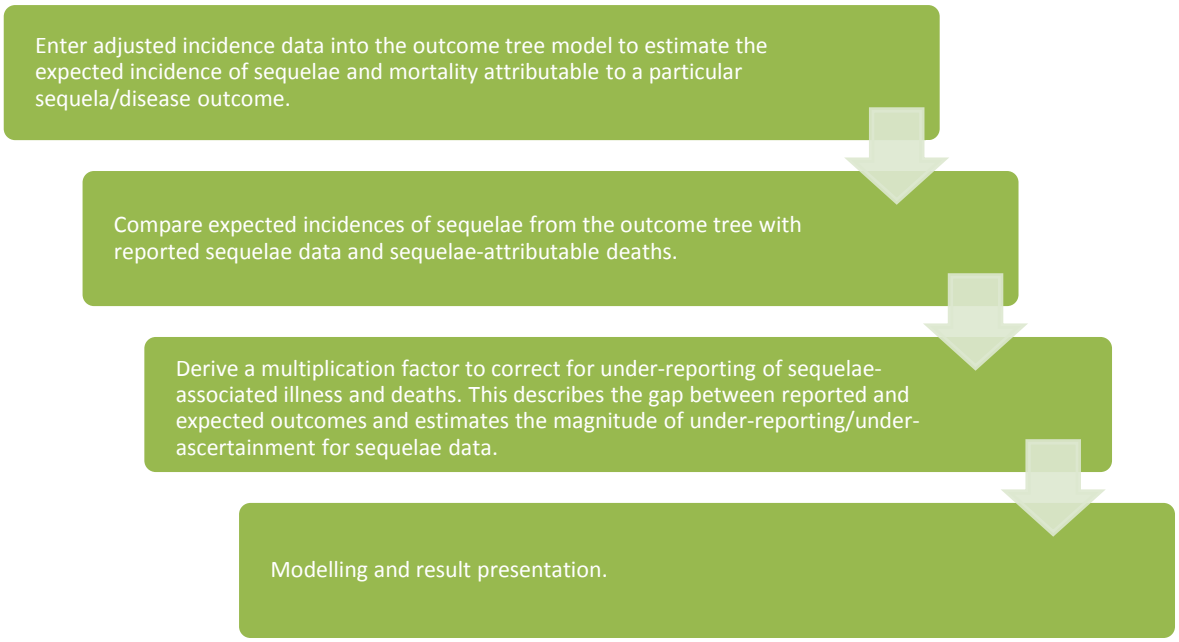


Longitudinal comparison (applicable for data on sequelae) (see Figure 6 below)

- Enter the adjusted incidence data (as per Figure 5) into the outcome tree model in order to estimate the expected incidence of sequelae and the mortality attributable to particular sequelae.
- Compare expected incidences of sequelae, as determined from the outcome tree, with reported data on sequelae and sequelae-attributable deaths.
- Based on this comparison, derive multiplication factors to correct for under-ascertainment of sequelae, associated morbidity and deaths.
  - Apply multiplication factors to correct for under-reporting and under-ascertainment.

    The observed differences between reported and expected outcomes will provide information on under-reporting and under-ascertainment and possibly validate the consistency of the outcome tree modelling framework.

Figure 6. Flowchart for longitudinal under-reporting (sequelae data)



Modelling the burden of a particular pathogen

Based on the defined outcome tree of a particular pathogen, a search is done for incidence data on all health states defined in the outcome tree.

If available raw incidence data (morbidity and mortality) is collected for each health state for a three-year period. These three years of crude incidence data are used as input for our model. Based on this period, a crude annual mean incidence can be estimated, stratified by age and sex. Where necessary these incidences have to be corrected by a factor α to correct for under-reporting, and by a factor β to correct for under-ascertainment. Both factors may be age-dependent and are specific to disease outcomes. These multiplication factors have to be obtained from the literature. For health outcomes with multiple sources the absolute number of cases, as obtained from disease-specific registers, will be multiplied by an attributable fraction δ obtained from the literature. For sequelae, but also for other health outcomes where no raw incidence data is available, we will estimate the number of cases by modelling the probability of occurrence of outcomes, taking into consideration the conditional dependency of the different health outcomes, as illustrated in the outcome tree.

The disease burden is calculated from the number of notified cases as follows. Let $n(a)$ be the number of notified cases at age a of a specific communicable disease with health outcomes $i=1,\ldots,k$. Let $1/\sigma(a)$ be the proportion of cases in the community that are notified. The factor $\sigma(a)$ is a product of factors α and β as defined above. Let $p_i(a)$ be the probability of developing outcome i for an individual infected at age a (possibly after transition through a chain of subsequent outcomes, i.e. p_i are not necessarily independent). Then $\sigma(a)n(a)$ is the number of all infections in the community at age a and $\sigma(a)n(a)p_i(a)$ is the number of cases of outcome i related to infections at age a. Let $\mu_i(a)$ be the probability that death occurs as a consequence of disease outcome i if infected at age a. Then $d_i(a)=\sigma(a)n(a)\,p_i(a)\,\mu_i(a)$ is the number of deaths due to disease outcome i. The $p_i(a)$ and $\mu_i(a)$ are parameters that have to be estimated from disease progression studies. The multiplication factor $\sigma(a)$ may be derived from various types of information including proportion of symptomatic cases, rate of diagnosis and notification rates. We can now integrate the outcome tree approach with the original DALY computation, as outlined on page 17.

Now we have $n_i(a)=\sigma(a)\,n(a)p_i(a)$ as the number of cases at age a of health states $i=1,\ldots,k$, where k is the number of different health outcomes resulting from the disease and $d_i(a)$, as defined above. Now as before, if $E(a)$ is the remaining life expectancy at age a, and a_{max} the maximum age, then

$$YLL=\sum_{a=0}^{a_{\max}}\sum_{i=1}^{k}d_i(a)E(a)=\sum_{a=0}^{a_{\max}}\sum_{i=1}^{k}\sigma(a)n(a)p_i(a)\mu_i(a)E(a)$$

The *YLD* is calculated as the product of the illness duration and the disability weight of a specific disease outcome, accumulated over the number of cases of all health outcomes. By denoting the duration of health outcome *i* as D_i and its disability weight as w_i, we obtain

$$YLD = \sum_{a=0}^{a_{\max}} \sum_{i=1}^{k} n_i(a) D_i w_i = \sum_{a=0}^{a_{\max}} \sum_{i=1}^{k} \sigma(a) n(a) p_i(a) D_i w_i$$

$E(a)$ and D_i have time as a unit, all other quantities are dimensionless numbers. Age is discretised into (a_{max} +1) yearly age classes.

By adding *YLL* and *YLD* we obtain DALYs (Disability-Adjusted Life Years) for pathogen *i*.

Software used for burden calculation

The disease burden models will be calculated in Microsoft Excel using @Risk, an add-in to Microsoft Excel that samples from probability distributions to obtain parameter values. Any uncertain parameter value in one of the Excel spreadsheet calculations can be replaced with an appropriated @Risk probability distribution function (e.g. Normal, Exponential, Uniform, etc.). The @Risk probability distribution function represents a range of possible values for a cell instead of limiting it to one value. Once all uncertain values have been replaced by appropriated @Risk functions and the output values defined, simulations can be run. @RISK recalculates the Excel spreadsheet model thousands of times, each time sampling random values from the @RISK functions entered, placing them in the model and recording the outcome. The simulations result in a sample of outcomes from a distribution of possible outcomes. In addition to expected outcomes, we therefore also obtain distributional properties of the outcome, such as the variance.

The application of more advanced software will be explored to incorporate temporal variation in incidence.

Modelling uncertainty

The data necessary for the quantitative estimates of disease burden are often limited and/or absent, resulting in a degree of uncertainty.

As a general rule, total uncertainty can be broken down into variability and uncertainty (Vose 2001). Variability is defined as the 'inherent heterogeneity of a system' (e.g. variations in the length of hospital stay for different patients). Uncertainty is usually defined as 'a lack of perfect knowledge about a factor in the model that represents the system' (more data, or better data would reduce uncertainty).

A further distinction of uncertainty would be:

- Statistical uncertainty due to small sample sizes
- Systematic uncertainty, due to the quality of data which are not fully representative of the country and/or disease under investigation
- Uncertainty due to lack of data.

For base case estimates, in our model we have chosen to use only point estimates of variable parameters. The impact of variability can be assessed during a second stage of the analysis using sensitivity analysis.

However, uncertainty will be explicitly modelled by incorporating probability distributions for uncertain parameters. These parameters will then be sampled from their probability distributions using Monte Carlo simulation to estimate predictive intervals.

Systematic uncertainty and uncertainty due to lack of data are best presented through scenario analysis, as will be the case in this project. In a scenario analysis only one parameter will be changed at a time.

Model outcome

The outcomes of the model are disease burden estimates or Disability-Adjusted Life Years (DALYs) per year associated with a particular infectious agent and related sequelae for a particular country. Data are presented in both aggregated form (DALYs per year) and in disaggregated form (YLD per year and YLL per year), from a societal perspective (DALYs/YLL/YLD per year per country and per 100 000 population per country) and from an individual perspective (DALYs/YLL/YLD per year per individual infected case). Where relevant in order to assess intervention impact, YLD and YLL will be analysed for major sequelae.

Where possible, the estimates obtained will be compared with estimates derived from other studies. This mainly relates to the GBD study conducted by WHO.

Results presentation

Mean estimates will be presented along with the associated statistical uncertainty. Scenario analysis is used to represent systematic uncertainty or uncertainty due to lack of data.

The data necessary for quantitative estimates of disease burden are often limited and/or absent, resulting in a certain degree of uncertainty that will be presented in the results.

Strengths and limitations

Disease definition

Disease definition is not always the same throughout Europe and this can lead to potential pitfalls when estimating disease burden.

Incidence data

When making the calculation the non-availability of incidence data, attributable fractions or conditional probabilities are potential limitations which, in some cases, might lead to the exclusion of a health outcome from the outcome tree for a selected pathogen. As a consequence the true disease burden may be underestimated for some of the pathogens selected.

Although the aim is to conduct disease burden calculations of the selected pathogens for all EU Member States and EEA/EFTA countries, non-availability of incidence data or poor data quality are potential reasons for excluding some of the selected pathogens in particular countries.

Age-specific data

The non-availability or poor data quality of age-specific data for both fatal outcomes and chronic health states will increase the uncertainty of the average disease burden estimate.

Duration of health states

For some health states information on duration might be limited, increasing the uncertainty of the disease burden estimate.

Time period

The aim of the BCoDE project is to obtain a disease burden estimate representing the years 2005-2007. However, the data available from these years might not always be complete or representative of the epidemiology for a disease. We will therefore also make use of community-cohort studies conducted in earlier years. However, factors such as demographic change, behavioural change in the population and the introduction of medical treatments might have had an impact on the dynamic of an infectious pathogen over time and the results of such studies should therefore be interpreted with care.

Incidental outbreaks of communicable diseases during the period under consideration lead to an overestimation (underestimation) of the disease burden if not noted and corrected.

Under-reporting/under-ascertainment

For most of the selected pathogens, some corrections will have to be made for under-reporting. However, these are only estimations of the true burden in the community and should therefore be interpreted with care.

Steady-state assumption

Medical interventions and therapies, demographic changes, climate change and other environmental and social changes have influenced and continue to influence the epidemiology of infectious diseases. These factors determine the transmission dynamics of infectious diseases and therefore influence the time trends of incidence and prevalence. Assuming a steady state limits the ability of the modelling approach to project disease burden into the future.

The methodological approach described in the current protocol estimates the disease burden in a situation where the above factors are assumed to be unchanging over a long period of time (known as the steady state

assumption). When projecting estimates into the future or assessing the impact of interventions in the past, temporal changes in environment and demography should be taken into account. At a later stage we therefore plan to use mathematical methods to model the potential dynamics of infectious agents over time.

Antibiotic resistance

Information on the incidence of infections with resistant pathogen strains is limited. For most pathogens, information on resistant strains is mainly collected in hospitals and cannot be extrapolated to the community. In principle, burden could be estimated by assessing the impact of resistance on various parts of the outcome tree. An infection with a resistant strain may have different transition probabilities and durations of sequelae compared to sensitive strains. However, infections with resistant strains are often concentrated in populations with other underlying conditions or immuno-compromised patients. Using our methodology it will be possible to conduct scenario analysis to assess the likely impact of resistance on disease burden.

Comparability with GDB estimates

Estimates derived with the methods outlined in this document will be comparable with GBD estimates to a certain extent. In particular, they will be comparable to estimates computed without discounting and age weighting for those diseases or groups of diseases that are included in both studies. Our estimates will differ from GBD estimates in that they will attribute certain sequelae to infectious causes.

A question for comparison will be whether total burden estimates derived with our methods are higher or lower than GBD estimates, or whether our methods only lead to a different distribution of burden across disease classes.

Next steps

Using the methods described in this document disease burden estimates will be obtained that enable the impact of communicable diseases on population health to be compared between communicable diseases and with other non-communicable diseases and conditions. These comparisons can be made at the national level and between countries. Estimates of the burden of communicable diseases provide a basis for prioritising health policy and research funds.

This protocol represents a first step towards delineating a concise description of the methodology used to estimate the burden of communicable diseases. As each step is being reviewed during the ongoing test phase of Work Package 2, the protocol will be updated to reflect experience with the practical implementation.

Annex 1. Diseases (pathogens and syndromes)

Respiratory tract infections
Seasonal influenza
Legionellosis
Tuberculosis
STI, including HIV and blood-borne viruses
Chlamydia
Gonococcal infections
Hepatitis B
Hepatitis C
HIV
Syphilis
Food-and water-borne diseases
Campylobacteriosis
Cryptosporidiosis
Infection with VTEC/STEC
Giardiasis
Hepatitis A
Listeriosis
Salmonellosis
Shigellosis
Toxoplasmosis
Variant Creutzfeldt-Jakob disease
Emerging and vector- borne diseases
Q fever
Tick-borne encephalitis
Vaccine-preventable diseases
Diphtheria
Invasive Haemophilus influenza disease
Invasive pneumococcal infections
Measles
Invasive meningococcal disease
Mumps
Pertussis
Poliomyelitis
Rabies
Rubella
Tetanus

Nosocomial infections (healthcare-associated infections)
Urinary tract infections (URI)
Surgical site infections (SSI)
Lower respiratory tract infections (LRTI)
Blood stream infections (BSI)
Gastro-intestinal Infections (GII) Clostridium difficile
Skin and soft tissue infections (SST)
Diseases as a consequence of infection
Cervical cancer (due to HPV)
Gastric cancer (due to H. pylori)
End stage liver diseases (due to HBV and HCV)
Primary liver cancer (due to HBV and HCV)

Annex 2. Criteria used when selecting pathogens/diseases

The master list is based on the objectives set out in the BCoDE Call for Proposal and includes:

- Forty-nine communicable diseases and special health issues specified under Decision 2119/98/EC with amendments.
- Other diseases and conditions which may result from infections, including infection-associated cancers. Based on a strong recommendation by the experts assembled at the BCoDE expert workshop in March 2010, the following conditions will be included in this category: end-stage liver disease or primary liver cancer related to hepatitis B virus (HBV), or hepatitis C virus (HCV) infection; cervical cancer related to human papilloma virus infection and gastric cancer related to *Helicobacter pylori* infection.
- Other conditions were added following discussions with ECDC Disease-Specific Programmes (DSP) and include diseases such as leishmaniasis, Lyme disease, tick-borne encephalitis and viral haemorrhagic fevers (including Crimean-Congo haemorrhagic fever, diseases caused by hantaviruses, dengue, Rift Valley fever, diseases caused by the Marburg filovirus and chikungunya).

Exclusion criteria:

1. Expert consensus (based on expert consideration of published evidence) that the burden of a disease/special health issue is negligible, mostly due to very low incidence, defined as equal to or lower than 1/100 000 per year.
2. Extreme methodological and practical difficulties foreseen in estimating the burden due to:
   - Lack of data and very low probability of data generation over the next five years
   - Serious methodological difficulties foreseen, for example in defining a disease and/or special health condition
   - Lack of disability weights for the most important health states of outcome trees, and absence of similar alternatives.

Exemptions:

If Criterion 1 for exclusion is met but:

- the negligible burden is believed to be low due to vaccination programmes with high coverage, or
- there is an outbreak potential based on history or theoretical considerations,

then the disease/special health issue will be kept on the final list and an attempt made to estimate the burden in the absence of intervention (e.g. by using the prevented fraction estimates in the case of interventions and mathematical modelling in the case of outbreak potential).

Process for defining the final list:

- The master list will be reviewed by the ad hoc working group of the ECDC Advisory Forum.
- Decisions will be based on a simple majority of responses from the group.
- The working group will start by considering the methodological issues (Criterion 2 above, Column A in the example table below). If the majority of responses relating to the feasibility of burden estimation is negative, the disease/health issue will be dropped from the list.
- The resulting shortlist will be assessed against the incidence (Criterion B: Above very low relative incidence i.e. $\geq$1/100 000 per year). Diseases that score a majority of "No" answers under Criterion B need to have a majority of "Yes" under Criterion C (Outbreak potential) or D (Vaccine-preventable with widely used vaccine) in order to be kept on the list.
- The final list will then be discussed with the Chief Scientist of ECDC and the management team of the project consortium.

Where necessary, experts can add or remove other diseases, by producing thorough justification which must be approved by the Chief Scientist of ECDC.

Table 2. Decision tool to facilitate selection of diseases/health conditions

Diseases/health issues		Criterion A	Criterion B	Criterion C	Criterion D
Respiratory tract infections		Feasibility based on data availability or methodological issues Y/N	Above very low relative incidence (i.e. ≥1/100,000 per year) Y/N	Outbreak potential? Y/N	Vaccine-preventable with widely used vaccine? Y/N
1	Seasonal influenza				
2	Avian influenza				
3	Legionellosis				
4	Tuberculosis				
5	etc.				

Annex 3. Comparison of pathogen-based DALY approach with incident case modelling approach to generate DALYs

Table 3. Comparison of pathogen-based DALY methodology and incident case modelling approach

	Extent of complexity and data requirement	Inclusion of burden due to long-term sequelae	Population base	Consideration of time-to-event	Consideration of competing risks and background mortality	Compatibility of estimate with data from GBD project	Applicability of approach
Pathogen-based DALY approach	+++	Yes	Cross-sectional, multi-cohort	No	No	?	Prioritisation + comparison
Incident cases modelling approach*	++++	Yes	Longitudinal, multi- or single-cohort	Yes	Yes	No (see time-to-event and competing event)	Prioritisation + comparison + forecasting + economic evaluation

** Suitable for both DALY and QALY calculations. This approach focuses exclusively on incident disease cases. The burden of disease is derived by modelling the fate of these cases over their remaining lifetime and comparing the modelled life-expectancy or quality-adjusted life-expectancy with the respective expectancies of a standard population without the disease. Longitudinal, single- or multi-cohort consideration (closed or open cohort) requires a valid natural history model. Modelling outcome: incremental QALY or DALY.*

References

1. Murray CJ, Lopez AD. Quantifying disability: data, methods and results. Bull World Health Organ. 1994;72(3):481-94.
2. Murray CJL, Lopez AD. Rethinking DALYs. The Global Burden of Disease: a comprehensive assessment of mortality and disability from diseases, injuries and risk factors in 1990 and projected to 2020. World Health Organization: Geneva 1996 1-98.
3. Murray CJL, Lopez AD, (eds.). Global health statistics: a compendium of incidence, prevalence and mortality estimates for over 200 conditions (Global burden of disease and injuries series; II). Cambridge: Harvard School of Public Health on behalf of the World Health Organization and the World Bank; 1996.
4. Murray CJL, Acharya AK. Understanding DALYs. Journal of Health Economics. 1997;16(6):703-30.
5. van Lier A, Havelaar AH. Disease burden of infectious disease: a pilot study. Bilthoven: National Institute of Public Health and the Environment 2007. Rpt No.: 215011001.
6. Mitchell AE, Sivitz LB, Black RE. Gulf war and health. Volume 5: Infectious Diseases. Washington, D.C.: The National Academies Press; 2007.
7. Schwarzinger M, Stouthard ME, Burstrom K, Nord E. Cross-national agreement on disability weights: the European Disability Weights Project. Popul Health Metr. 2003 Nov 21;1(1):9.
8. Johnson JA, Luo N, Shaw JW, Kind P, Coons SJ. Valuations of EQ-5D health states: are the United States and United Kingdom different? Med Care. 2005 Mar;43(3):221-8.
9. Torrance GW. Measurement of health state utilities for economic appraisal. J Health Econ 1986;5(1):1-30
10. Field MJ, Gold MR. Summarizing population health directions for the development and application of population metrics. Washington, D.C.: National Academy Press;1998.
11. Essink-Bot ML, Bonsel GJ. How to derive disability weights. In: Murray CJL, Salomon JA, Mathers CD, Lopez A, World Health Organization (eds.) Summary measures of population health: Concepts, ethics, measurement and applications. Geneva: World Health Organization; 2002;449-65.
12. Gold MR, Stevenson D, Fryback DG. HALYS and QALYS and DALYS, Oh my: Similarities and differences in summary measures of population health. Annu Rev Public Health. 2002;23:115-34.
13. Torrance GW. Social preferences for health states. An empirical evaluation of three measurement techniques. Socio-Economic Planning Sciences 1976;10(3):129-136.
14. Nord E. My goodness - and yours: a history and some possible futures of DALY - meanings and valuation procedures. In: Murray CJL, Salomon JA, Mathers CD, Lopez A (eds). Summary measures of population health: concepts, ethics, measurement and applications. Geneva: World Health Organization 2002;139-46.
15. Health Protection Agency (HPA) UK. Health protection in the 21st century - Understanding the burden of disease; preparing for the future. Available from: (http://www.hpa.org.uk/webw/HPAweb&HPAwebStandard/HPAweb_C/1203064747564?p=1153846674362). London 2005.
16. Mead PS, Slutsker L, Dietz V, McCaig LF, Bresee JS, Shapiro C, et al. Food-related illness and death in the United States. Emerg Infect Dis. 1999 Sep-Oct;5(5):607-25.
17. Mead PS, Slutsker L, Griffin PM, Tauxe RV. Food-related illness and death in the United States. Reply to Dr. Hedberg. Emerg Infect Dis. 1999 Nov;5(6):841-2.
18. Van Damme P, Tormans G, Beutels P, Van Doorslaer E. Hepatitis B prevention in Europe: a preliminary economic evaluation. Vaccine. 1995;13 (Suppl 1):S54-7.
19. Cleaveland S, Fèvre EM, Kaare M, Coleman PG. Estimating human rabies mortality in the United Republic of Tanzania from dog bite injuries. Bulletin of the World Health Organization. 2002;80:304-10.
20. Odiit M, Coleman PG, Liu WC, McDermott JJ, Fèvre EM, Welburn SC, et al. Quantifying the level of under-detection of Trypanosoma brucei rhodesiense sleeping sickness cases. Tropical Medicine & International Health. 2005;10(9):840-9.
21. Fallah M, Kharazm. E. Correction for under-ascertainment in cancer cases in the very elderly (aged 75+): external reference method. Cancer Causes and Control. 2008;19(7).
22. Fallah M, Kharazm. E. Iran cancer incidence should be corrected for under-ascertainment in cancer cases in the elderly (aged 65+). Asian Pacific J Cancer Prev. 2007;8:348-52.
23. Majowicz SE, Musto J, Scallan E, Angulo FJ, Kirk M, O'Brien SJ, et al. The global burden of nontyphoidal Salmonella gastroenteritis. Clin Infect Dis. 2010 Mar 15;50(6):882-9.

24. Crowcroft NS, Andrews N, Rooney C, Brisson M, Miller E. Deaths from pertussis are underestimated in England. Archives of Disease in Childhood. 2002 May 2002;86(5):336-8.

25. Henson SJ, Majowicz SE, Masakure O, Sockett PN, MacDougall L, Edge VL, et al. Estimation of the costs of acute gastrointestinal illness in British Columbia, Canada. International Journal of Food Microbiology. 2008;127(1-2):43-52.

26. Majowicz SE, Dore K, Flint JA, Edge VL, Read S, Buffett MC, et al. Magnitude and distribution of acute, self-reported gastrointestinal illness in a Canadian community. Epidemiol Infect. 2004 Aug;132(4):607-17.

27. Thomas MK, Majowicz SE, Pollari F, Sockett PN. Burden of acute gastrointestinal illness in Canada, 1999-2007: Interim summary of NSAGI activities. Can Commun Dis Rep. 2008 May;34(5):8-15.

28. Szucs TD, Müller D. Influenza vaccination coverage rates in five European countries-a population-based cross-sectional analysis of two consecutive influenza seasons. Vaccine. 2005;23(43):5055-63.

29. Grulich AE, Visser RO, Smith AMA, Risse CE, Richters J. Sex in Australia: Sexually transmissible infection and blood-borne virus history in a representative sample of adults. Australian and New Zealand Journal of Public Health. 2003;27(2):234-41.

30. Imhoff B, Morse D, Shiferaw B, Hawkins M, Vugia D, Lancea-Parker S, et al. Burden of self-reported acute diarrheal illness in FoodNet surveillance areas, 1998-1999. Clinical Infectious Diseases. 2004;38(s3):S219-S26.

31. Fitzgerald M, Scallan E, Collins C, Crowley D, Daly L, Devine M, et al. Results of the first population-based telephone survey of acute gastroenteritis in Northern Ireland and the Republic of Ireland. Euro Surveill. 2004;8(18).

32. Wheeler JG, Sethi D, Cowden JM, Wall PG, Rodrigues LC, Tompkins DS, et al. Study of infectious intestinal disease in England: rates in the community, presenting to general practice, and reported to national surveillance. BMJ. 1999 April 17, 1999;318(7190):1046-50.

33. Sethi D, Wheeler J, Rodrigues L, Fox S, Roderick P. Investigation of under-ascertainment in epidemiological studies based in general practice. Int J Epidemiol. 1999 February 1, 1999;28(1):106-12.

34. de Wit MAS, Koopmans MP, Kortbeek LM, van Leeuwen NJ, Bartelds AI, van Duynhoven YT. Gastroenteritis in sentinel general practices,The Netherlands. Emerg Infect Dis 2001;7(1):82-91.

35. de Wit MAS, Koopmans MPG, Kortbeek LM, Wannet WJB, Vinje J, van Leusden F, et al. Sensor, a population-based cohort study on gastroenteritis in the Netherlands: incidence and etiology. Am J Epidemiol. 2001 October 1, 2001;154(7):666-74.

36. Parish WL, Laumann EO, Cohen MS, Pan S, Zheng H, Hoffman I, et al. Population-based study of chlamydial infection in China: A hidden epidemic. JAMA. 2003 March 12, 2003;289(10):1265-73.

37. Klavs I, Rodrigues LC, Wellings K, Weiss HA, Hayes R. Sexual behaviour and HIV/sexually transmitted infection risk behaviours in the general population of Slovenia, a low HIV-prevalence country in central Europe. Sexually Transmitted Infections. 2009 April 2009;85(2):132-8.

38. Hsieh YH, Kuo MJ, Hsieh TC, Lee HC. Under-reporting and under-estimation of gonorrhea cases in the Taiwan National Gonorrhea Notifiable Disease System in the Tainan region: evaluation by a pilot physician-based sentinel surveillance on Neisseria gonorrhoeae infection. International Journal of Infectious Diseases. 2009;13(6):e413-e419.

39. de Melker HE, Versteegh FGA, Schellekens JFP, Teunis PFM, Kretzschmar M. The incidence of Bordetella pertussis infections estimated in the population from a combination of serological surveys. Journal of Infection. 2006;53(2):106-13.

Bibliography

Anand S, Hanson K. DALYs: efficiency versus equity. World Development. 1998;26(2):307.

Anand S, Hanson K. Disability-adjusted life years: a critical review. Journal of Health Economics. 1997;16(6):685.

Arnesen T, Kapiriri. L. Can the value choices in DALYs influence global priority-setting? Health policy (Amsterdam, Netherlands). 2004;70(2):137-49.

Barendregt JJ, Bonneux L, Van der Maas PJ. DALYs: the age-weights on balance. Bull World Health Organ. 1996;74(4):439-43.

Batz M, Mangen M-J, Käsbohrer A, Hald T, Morris JGJ, Taylor M, et al. Priority setting for food-borne and zoonotic pathogens - A summary for an international conference organized jointly by the EU MED-VET-NET Network of Excellence and the US Food Safety Research Consortium: MED-VET-NET2007 Aug 2007. Rpt No.: 07-001.

Baud D, Peter O, Langel C, Regan L, Greub G. Seroprevalence of Coxiella burnetii and Brucella abortus among pregnant women. Clinical Microbiology and Infection. 2009;15(5):499-501.

Begg S, Voss T, Barker B, Stevenson C, Stanley L et al. Burden of disease and injury, Australia 2003. Canberra: Australian Institute of Health and Welfare. 2007. Available from: http://catalogue.nla.gov.au/Record/4085387

Bradshaw D, Groenewald P, Laubscher R, Nannan N, Nojilana B, Norman R, et al. Initial burden of disease estimates for South Africa, 2000. S Afr Med J. 2003 Sep;93(9):682-8.

Brazier JE, Ratcliffe J, Salomon JA, Tsychiya A. Measuring and valuing health benefits for economic evaluation. New York: Oxford University Press Inc.; 2007.

Chapman G. The burden of disease in Zimbabwe in 1997 as measured by Disability-Adjusted Life Years lost. Tropical Medicine & International Health. 2006;11(5):660.

de Greeff SC, Spanjaard L, Dankert J, Hoebe CJ, Nagelkerke N, de Melker HE. Under-reporting of meningococcal disease incidence in the Netherlands: results from a capture-recapture analysis based on three registration sources with correction for false positive diagnoses. Eur J Epidemiol. 2006;21(4):315-21.

Dodhia H, Phillips K. Measuring burden of disease in two inner London boroughs using Disability-Adjusted Life Years. J Public Health 2008 Apr 8.

Drummond MF, O'Brien. B, Stoddart GL, Torrance, GW. Methods for the economic evaluation of health care programmes. Oxford: Oxford University Press; 1997.

ESEN2. European Sero-Epidemiology Network 2 Available from: http://www.hpa.nhs.uk/HPA/AboutTheHPA/WhatWeDo/InternationalWork/1204542875968/

Frost FJ, Craun G, Mihaly K, Gyorgy B, Calderon R, Muller T. Serological responses to Cryptosporidium antigens among women using riverbank-filtered water, conventionally filtered surface water and groundwater in Hungary. Journal of Water and Health. 2005;03.01:77-82.

Frost FJ, Fea E, Gilli G, Biorci F, Muller TM, Craun GF, et al. Serological evidence of Cryptosporidium infections in southern Europe. European Journal of Epidemiology. 2000;16(4):385-90.

Fundação Nacional de Saúde - Ministério da Saúde de Brazil 2002. Relatorio final do projecto estimativa da carga de doenca do Brasil 1998.

Fundação Nacional de Saúde - Ministério da Saúde de Brazil. Relatorio final do projecto estimativa da carga de doenca do Brasil 1998. 2002.

Gadelha AMJ, Leite IdC, Valente JG, Schramm JMdAS, Profa Portela MC, Campos MR. Relatorio final do projeto estimativa da carga de doenca do Brasil - 1998 2002. Publ. 2002.

Haagsma JA, Havelaar AH, Janssen BM, Bonsel GJ. Disability-Adjusted Life Years and minimal disease: application of a preference-based relevance criterion to rank enteric pathogens. Popul Health Metr. 2008;6:7.

Hannu T, Mattila L, Rautelin H, Pelkonen P, Lahdenne P, Siitonen A, et al. Campylobacter-triggered reactive arthritis: a population-based study. Rheumatology. 2002 March 1, 2002;41(3):312-8.

Havelaar AH, Van Duynhoven YTHP, Nauta MJ, Bouwknegt M, Heuvelink AE, De Wit GA, et al. Disease burden in the Netherlands due to infections with Shiga toxin-producing Escherichia coli O157. Epidemiology and Infection. 2004;132(03):467-84.

Khealani BA, Wasay. M. The burden of stroke in Pakistan. International Journal of Stroke. 2008;3:293-296.

Kominski GF, Simon PA, Ho A, Luck J, Lim YW, Fielding JE. Assessing the burden of disease and injury in Los Angeles County using Disability-Adjusted Life Years. Public Health Rep. 2002 Mar-Apr;117(2):185-91.

Kretzschmar M, Teunis PF, Pebody RG. Incidence and reproduction numbers of pertussis: estimates from serological and social contact data in five European countries. PLoS Med. 2010 Jun;7(6):e1000291.

Krishnamoorthy K, Harichandrakumar KT, Krishna Kumari A, Das LK. Burden of chikungunya in India: estimates of Disability-Adjusted Life Years (DALY) lost in 2006 epidemic. J Vector Borne Dis. 2009 Mar;46(1):26-35.

Krupnick AJ. Valuing health outcomes: Policy choices and technical issues. Washington, D.C.,: Resources for the Future 2004.

Lai T, Habicht J. Kiivet RA. Measuring burden of disease in Estonia to support public health policy. Eur J Public Health. Oct 1, 2009;19(5):541-7.

Lopez AC, Mathers CD, Ezzati M, Jamison DT, Murray CJL, Disease Control Priorities Project. Global burden of disease and risk factors. New York, N.Y.: Washington, D.C.: Oxford University Press; World Bank; 2006.

Lopez AD, Mathers CD, Ezzati. M., Jamison. D.T., Murray CJL. Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. Lancet 2006;367(9524):1747.

Lozano R, Murray CJL, Frenk J, Bobadilla JL. Burden of disease assessment and health system reform: Results of a study in Mexico. Journal of International Development. 1995;7(3):555-63.

Luz PM, Grinsztejn B, Galvani AP. Disability-Adjusted Life Years Lost to dengue in Brazil. Tropical Medicine & International Health 2009;14(2):237-46.

Mahapatra P. Estimating national burden of disease: the burden of disease in Andhra Pradesh – 1990s. Hyderabad Institute of Health Systems. 2002.

Mangen MJ, Batz MB, Kasbohrer A, Hald T, Morris Jr JG, Taylor M, et al. Integrated approaches for the public health prioritization of food-borne and zoonotic pathogens. Risk Anal. 2010;30(5):782-97.

Mangen MJJ, Havelaar AH, Bernsen RAJAM, van Koningsveld R, De Wit GA. The costs of human Campylobacter infections and sequelae in the Netherlands: A DALY and cost-of-illness approach. Food Economics - Acta Agriculturae Scandinavica, Section C. 2005;2(1):35-51.

Mathers CD, Bernard C, Iburg KM, Inoue M, Ma Fat D, Shibuya K, et al. Global burden of disease in 2002: data sources, methods and results. Geneva: World Health Organization; 2003.

Mathers CD. Introduction to burden of disease. Burden of disease workshop; Bielefeld, Germany. January 2006.

McMahon BJ, Alward WL, Hall DB, Heyward WL, Bender TR, Francis DP, et al. Acute hepatitis B virus infection: relation of age to the clinical expression of disease and subsequent development of the carrier state. J Infect Dis. 1985 Apr;151(4):599-603.

Med-Vet-Net. Campylobactor serology. 2007.

Melse JM, Essink-Bot ML, Kramers PG, Hoeymans N. A national burden of disease calculation: Dutch disability-adjusted life-years. Dutch Burden of Disease Group. Am J Public Health. 2000 August 1, 2000;90(8):1241-7.

Michaud C, McKenna M, Begg S, Tomijima N, Majmudar M, Bulzacchelli M, et al. The burden of disease and injury in the United States 1996. Population Health Metrics. 1996;4(1):11.

Ministerio de Salud del Perú. Estudio de carga de enfermedad an el Peru 2004. Publ. 2006

Ministry of Health in Singapore. Singapore burden of disease study 2004. Publ. 2009

Ministry of Health, Refik Sysam Hygiene Center Presidency, School of Public Health: Başkent University. National burden of disease and cost effectiveness project, Burden of disease. Final report 2004.

Mitchell AE, Sivitz LB, Black RE. Gulf war and health. Volume 5: Infectious Diseases. Washington, D.C: The National Academies Press; 2007.

Muhlberger N, Schwarzer R, Lettmeier B, Sroczynski G, Zeuzem S, Siebert U. HCV-related burden of disease in Europe: a systematic assessment of incidence, prevalence, morbidity, and mortality. BMC Public Health. 2009;9:34.

Murray CJ. Quantifying the burden of disease: the technical basis for disability-adjusted life years. Bull World Health Organ. 1994;72(3):429-45.

Murray CJL, Acharya AK. Understanding DALYs. Journal of Health Economics. 1997;16(6):703-30.

Murray CJL, Lopez AD. Global mortality, disability, and the contribution of risk factors: Global Burden of Disease Study. Lancet. 1997 May 17;349(9063):1436-42.

Murray CJL, Lopez AD. Regional patterns of disability-free life expectancy and disability-adjusted life expectancy: Global Burden of Disease Study. Lancet. 1997;349(9062):1347-52.

Murray CJL, Lopez AD. The utility of DALYs for public health policy and research: a reply. Bulletin of the World Health Organization. 1997c;75(4):377-81.

Naghavi M, Abolhassani F, Pourmalek F, Lakeh M, Jafari N, Vaseghi S, et al. The burden of disease and injury in Iran 2003. Popul Health Metr. 2009;7:9.

Nardone A, Miller E. Serological surveillance of rubella in Europe: European Sero-Epidemiology Network (ESEN2). Euro Surveill 2004 2004;9:5-7.

Nojilana B, Norman R, Bradshaw D, van Stuijvenerg ME, Dhansay MA Labadorios D. Estimating the burden of disease attributable to vitamin A deficiency in South Africa in 2000. S Afr Med J. 2007:97(8 Pt2);748-53

Norman R, Mathee A, Barnes B, van der Merwe L, Bradshaw D. Estimating the burden of disease attributable to lead exposure in South Africa in 2000. S Afr Med J. 2007;97(8 Pt2);773-80

Siebert U. When should decision-analytic modelling be used in the economic evaluation of health care? The European Journal of Health Economics. 2003;4(3).

Simonsen J, Strid MA, Mølbak K, Krogfelt KA, Linneberg A, Teunis P. Sero-epidemiology as a tool to study the incidence of Salmonella infections in humans. Epidemiology and Infection. 2008;136(07):895-902.

Stouthard MEA, Essink-Bot ML, Bonsel GJ, Barendregt J, Kramers PGN, van de Water HPA, et al. Wegingsfactoren voor ziekten in Nederland. Amsterdam: Instituut voor Sociale Geneeskunde, Academisch Medisch Centrum Amsterdam; 1997.

Stouthard MEA, Essink-Bot M-L, Bonsel GJ, on behalf of the DDWG. Disability weights for diseases: A modified protocol and results for a western European region. Eur J Public Health. Mar 1, 2000;10(1):24-30.

Teunis PF, van der Heijden OG, de Melker HE, Schellekens JF, Versteegh FG, Kretzschmar ME. Kinetics of the IgG antibody response to pertussis toxin after infection with B. pertussis. Epidemiol Infect. 2002 Dec;129(3):479-89.

Torrance GW. Measurement of health state utilities for economic appraisal. J Health Econ. 1986 Mar;5(1):1-30.

Van Damme. P, Tormans G, Beutels P, Van Doorslaer, E. Hepatitis B prevention in Europe: a preliminary economic evaluation. Vaccine. 1995;13 (Suppl 1):S54-7.

van Hest NA, Smit F, Baars HW, De Vries G, De Haas PE, Westenend PJ, et al. Completeness of notification of tuberculosis in the Netherlands: how reliable is record-linkage and capture-recapture analysis? Epidemiol Infect. 2007 Aug;135(6):1021-9.

van Hest NA, Smit F, Verhave JP. Under-reporting of malaria incidence in the Netherlands: results from a capture-recapture study. Epidemiol Infect. 2002 Oct;129(2):371-7.

Vos T. The case against annual profiles for the valuation of disability weights. In: Murray CJL, Salomon JA, Mathers CD, Lopez AD, editors. Summary measures of population health: concepts ethics and applications. Geneva: World Health Organization; 2002:467-72.

World Health Organization. Global Burden of Disease (GBD) study. Operations manual: Final draft; January 20, 2009. Geneva: World Health Organization 2009.

World Health Organization. Global burden of disease: 2004 update. Geneva: World Health Organization 2008.

Zhang Y, Bi P, Hiller JE. Climate change and Disability-Adjusted Life Years. Journal of Environmental Health. 2007;70(3):32-6