European Commission

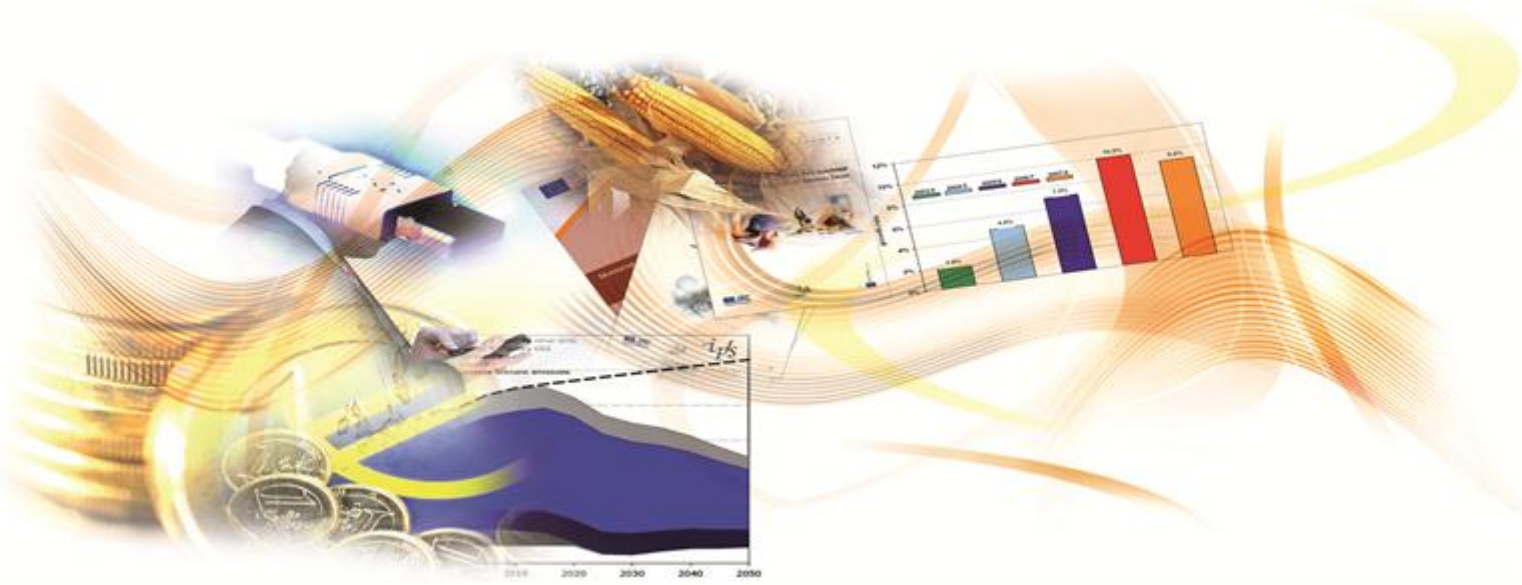
JRC SCIENCE AND POLICY REPORTS

Monitoring and Assessment Framework for the European Innovation Partnership on Active and Healthy Ageing (MAFEIP)

Conceptual description of the Monitoring and Assessment Framework for the EIP on AHA

Christian Boehler
Fabienne Abadie

2015



Report EUR 27412 EN

Joint Research Centre

European Commission
Joint Research Centre
Institute for Prospective Technological Studies

Contact information
Address: Edificio Expo. c/ Inca Garcilaso, 3. E-41092 Seville (Spain)
E-mail: jrc-ipts-secretariat@ec.europa.eu
Tel.: +34 954488318
Fax: +34 954488300

https://ec.europa.eu/jrc
https://ec.europa.eu/jrc/en/institutes/ipts

JRC96205

EUR 27412 EN

ISBN 978-92-79-50574-4 (PDF)

ISSN 1831-9424 (online)

doi:10.2791/290381

Luxembourg: Publications Office of the European Union, 2015

Abstract
After having identified a short list of candidate indicators for assessing the impact of the European Innovation Partnership on Active and Healthy Ageing (EIP on AHA) in the first and second reports on outcome indicators for MAFEIP, the next step in this project was to develop a quantitative approach that could be suited to establishing a link between candidate indicators and the EIP on AHA objectives. This report therefore conceptualises a model for estimating the impact of the Partnership's activities on its targets for health and sustainability of health and care system using the outcome indicators that were previously identified. In accordance with the EIP on AHA headline target of increasing the average healthy life expectancy of European citizens by two years by 2020, we took the methods to calculate Healthy Life Years (HLY) as a starting point, but adapted them to better accommodate the needs of MAFEIP. The rationale for this adaptation was to ensure the resulting model can adequately estimate the health impacts achieved by EIP on AHA commitments, and also to utilise data on indicators that are most frequently reported across EIP on AHA participants. The resulting model is based on a Markov process with three generic health states ('baseline health', 'deteriorated health' and 'death'), which can draw upon data from primary and secondary outcome indicators across populations, interventions, commitments and geographic domains. We discuss how the model's flexibility that allows it to be applied to different contexts could be enhanced further through the optional inclusion of additional health states or extensions for incorporating additional secondary indicators. We also discuss how to use the model for estimating the impact of activities delivered within the EIP on AHA on the sustainability of health and care systems in terms of the incremental impact of the interventions on health and care expenditure. We propose that the model should be implemented as a web-based monitoring tool to enable stakeholders within commitments to independently assess the impact of their respective interventions on health and sustainability of health and care systems, with the support and guidance of IPTS.

Table of Contents

List of Tables

List of figures

1. Introduction

1.1 Background

IPTS in cooperation with DG CNECT and DG SANCO is developing a monitoring framework to assess the evolution and impact of the European Innovation Partnership on Active and Healthy Ageing (EIP on AHA), through the "Monitoring and Assessment Framework for the EIP on AHA" (MAFEIP) project. The main objective of MAFEIP is to define a common monitoring framework, which should facilitate and harmonise the monitoring of the process of the EIP on AHA and of the outcome and output of the Action Groups within the EIP on AHA (not the individual commitments[1] to the six specific Action Groups). It will also seek to establish a link between the monitoring results and the EIP on AHA objectives, namely the Triple Win (quality of life; sustainability of health and care systems; and, innovation and growth) and the overall objective of increasing by two additional healthy life years the average healthy life span of European citizens by 2020.

1.1.1 The development process

The input of both the Action Group members and experts has been the basis for the initial monitoring framework developed by the EC. Experts were asked to provide input on the factors influencing quality of life and on WHO, OECD and Eurostat data (meeting June 2012 – August 2012). In addition, in August 2012 the Action Group members were asked to provide input about the methodology and indicators they are using and the kind of data they gather. After intensive interaction between the Action Group members and the experts and following the Action Group meetings on the further fine-tuning of the framework (September-October 2012), a meeting took place with the experts and two members of each Action Group to discuss remaining issues in November 2012 and to define the initial Monitoring Framework.

1.1.2 The initial monitoring framework

The objective of the outcome indicators is to monitor the factors influencing the Triple Win, for example:

- For Quality of life of patients/users, this may include changes in risk factors, nutrition or physical activity
- For Sustainability of health and care systems, changes in hospital admissions, or a shift towards home care instead of institutionalisation
- And for Innovation and growth possibilities, new employment opportunities.

The proposed monitoring framework consists of three columns, which represent the Triple Win as shown in Figure 1 below:

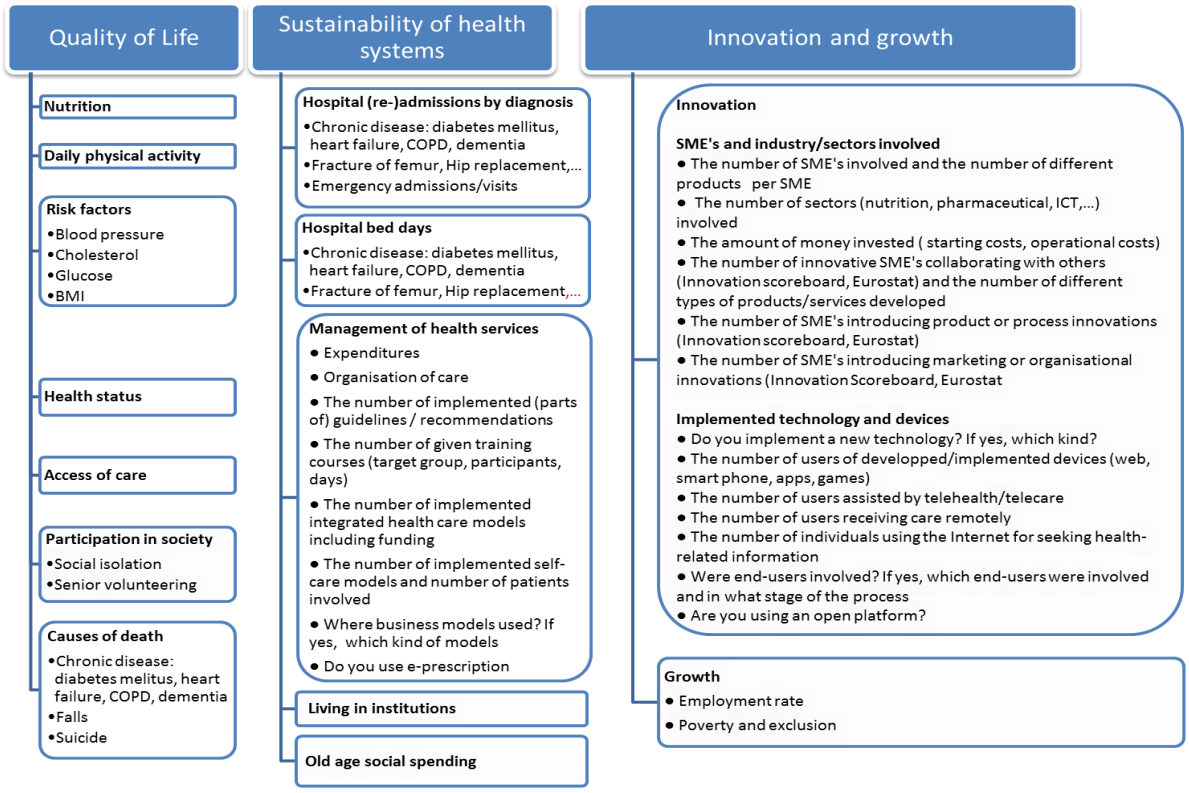
- Quality of Life
- Sustainability of health and care systems and
- Innovation and growth.

The multidimensional approach of the monitoring framework for the EIP on AHA entails a set of building blocks which contain different indicators. The monitoring framework and its three columns and building block system were originally devised in such a way that not all blocks of indicators are necessarily relevant for all Action Groups. Therefore, not all individual actions may contribute to all building blocks. However, each individual action should contribute to at least one building block. As a result, the outcome indicators in Figure 1 may constitute a long-list of potential indicators which

[1] Commitments are stakeholders who have committed to engage through a project or an initiative, and contribute to the objectives of an EIP on AHA action, or a group of actions.

may not always be equally fit for linking to the overarching objective of the EIP on AHA and the Triple Win, either because the Action Groups' commitments cannot be forced to use specific indicators or because some of the indicators initially thought of may not be adequate for estimating the impacts on the EIP on AHA higher objectives.

Figure 1: EIP on AHA initial Monitoring Framework (5th of November 2012)



1.2 Refining and operationalising the initial framework

In order to further develop and operationalise the EIP on AHA monitoring and assessment framework, the following exercises were carried out within the MAFEIP project (see MAFEIP Inception Report annex[2] and MAFEIP's first report on outcome indicators[3]:

I. A double-check of the interactive process with experts and Action Group members that led to the initial framework structure, including the review of the indicators mentioned in the survey of Action Group participants from August 2012. This interactive process led to the definition of the initial Monitoring Framework of the EIP on AHA (**Figure 1**).

II. A review of Reference Sites' good practices, which provided the EIP on AHA with cases of comprehensive, innovation-based approaches to active and healthy ageing. This review identified outcome indicators from a self-assessment questionnaire submitted by a total of 71 good practices in April 2013. Some of the indicators identified in the analysis were found to be used by a number of Reference Sites, others were very specific to a few or even single sites. General categories of impact were thus defined, under which specific indicators were further specified and the frequency of occurrence

2 http://publications.jrc.ec.europa.eu/repository/bitstream/JRC85879/jrc85879.pdf
3 http://publications.jrc.ec.europa.eu/repository/bitstream/JRC91162/jrc91162.pdf

across the Reference Sites was recorded. This led to the identification and categorisation of indicators, and the assessment of their relevance across reference sites.

III. A clustering exercise which consisted of a review of commitments' good practices for Action Groups A1, A3, B3 and D4 in order to identify the outcome indicators used by these good practices, after grouping the interventions described into clusters.

IV. A systematic review of the scientific literature in each of the six Action Group domains. Though the review was done through systematic searches in specialised scientific databases, it could not be exhaustive because of the limited time and resources available and the breadth of the fields concerned. However it provided some useful indications of whether targeted initiatives like those undertaken by EIP on AHA commitments lead to scientifically-sound outcomes.

V. A review of macro-level and intermediary-level indicators with a view to assessing the possibility of establishing the link between the outcomes of Action Groups' commitments and the higher objectives of the EIP on AHA.

VI. An analysis of data from a survey conducted across Action Groups to learn about the actual outcome indicators being used to measure the impact of commitments on the health and sustainability targets of the EIP on AHA.

To inform the choice of outcome indicators proposed for use in the MAFEIP project, the above data sources I to V were studied in conjunction. The rationale was that each single source of information may not be adequate to inform the choice of outcome indicators for the monitoring framework of the EIP on AHA on its own, however, together, the various sources may support or disqualify the use of particular outcome indicators. In addition, whilst not conceptualising quantitative models as this is the purpose of the current report, we took into account aspects of linking outcome indicators to the 2HLY and the Triple Win in order to discriminate between a set of "***primary outcome indicators***" and "***secondary outcome indicators***". Primary outcome indicators are not just relevant across a number of commitments and Action Groups, but also particularly well suited to establishing a quantitative link to the Triple Win and the headline target. Secondary outcome indicators may require more elaborate approaches to establishing a quantitative link to primary outcome indicators, and ultimately to the Triple Win and the headline target. Secondary indicators could either be relevant for several Action Groups (common secondary indicators), or targeted at particular Action Groups (specific secondary indicators).

The joint analysis of data sources I to V, combined with initial considerations regarding the ability of outcome indicators to be linked to the EIP on AHA objectives, led to the definition of a short list of potential outcome indicators for the quantitative monitoring of the Partnerships' impact on quality of life and the sustainability of health and care systems. The results of this exercise were reported in MAFEIP's first report on outcome indicators[4]. Subsequently, we used data from the survey conducted across Action Groups (data source VI) to confirm and further refine this shortlist of outcome indicators, and the results from this exercise have been reported in the "MAFEIP Second report on outcome indicators"[5]. Before describing in more detail the aim of this document, which is to propose a quantitative modelling strategy for MAFEIP that links health impact and health and care system outcomes to EIP on AHA objectives, we will briefly summarise below the results from MAFEIP's first and second reports on outcome indicators.

[4] See footnote 3

[5] http://publications.jrc.ec.europa.eu/repository/bitstream/JRC93431/jrc93431.pdf

1.3 Short-list of candidate indicators for quantifying EIP on AHA outcomes

As mentioned, the activities described above resulted in a short list of potential outcome indicators for quantitatively modelling the impact of EIP of AHA activities. The results are presented separately for the columns 'Quality of Life' and 'Sustainability of health and care systems'.

1.3.1 "Quality of Life" Column

Our provisional proposal for outcome indicators in the Quality of Life column of the monitoring framework is summarised in Table 1.

Table 1: Provisional proposal for outcome indicators within the QoL column

	Primary indicators	**Secondary indicators**
Common indicators	HRQoL Mortality	Risk Factors Physical Activity
Specific indicators	n.a.	Adherence to treatment Falls Frailty Nutrition Mental health / depression Cognitive decline Functional status

Key (primary) outcome indicators should be general enough to be relevant across all Action Groups of the EIP on AHA, but also sufficiently specific and sensitive to capture the impact of particular interventions within each Action Group. We identified **measures of health-related quality of life (HRQoL)**, in particular instruments suited to constructing a single metric that combines the valuation of a certain health state with the time spent in this state, like the EQ-5D-instrument; or instruments that could potentially be mapped towards the EQ-5D including - but not limited to - members of the SF-family (e.g. SF-36 or SF-6D). Accordingly, we also identified **mortality rates** as primary indicators, stratified by major target diseases such as heart failure, stroke, COPD, dementia, falls, or suicide as applicable to each Action Group. The analysis of survey data reported in the MAFEIP's second report on outcome indicators provided very strong support for the use of primary outcome indicators for the purposes of MAFEIP, with 63% of respondents reporting the use of HRQoL and/or mortality indicators.

Secondary outcome indicators generally require a more elaborate approach to linking outcomes to the EIP on AHA objectives and the headline target and they could either be relevant for several Action Groups (common secondary indicators), or targeted at particular Action Groups (specific secondary indicators). The **common** secondary indicator '**Risk factors'** includes routine clinical measures such as blood pressure, cholesterol levels, blood glucose, and body mass index. A change in risk factors due to interventions delivered within different commitments may result in a change in quality-adjusted life expectancy. A quantitative link could be established, for instance, through the use of validated risk stratification tools such as the Framingham risk equation [D'Agostino et al., 2008] or the QRISK-tool [Hippisley-Cox J et al., 2007] for estimating cardiovascular disease risk. The **common** secondary indicator **Physical Activity** initially appeared to be relevant across a number of commitments and Action Groups. In addition, some evidence could be found in the literature and studies exist that provide a link between changes in physical activity behaviour and respective changes in quality-adjusted life expectancy using Decision Analytic Modelling (DAM) [NICE, 2006: Anokye et al., 2011; Anokye et al., 2013]. However, the survey data analysed in MAFEIP's second report on outcome indicators showed that coverage through the inclusion of common secondary indicators (both risk factors and physical activity) could only be increased by 9%, i.e. from 63% to 72% of respondents to the survey.

Next to those secondary indicators that are commonly supported by commitments across Action Groups, we identified some additional candidate indicators after a review of the good practices presented by each Action Group in their respective booklets[6]. We refer to these indicators as specific secondary indicators as they are not generally applicable across all Action Groups, but rather relevant for one or a few Action Groups within the EIP on AHA only. Whilst these indicators were initially supported through a combined analysis of good practices and clustering data in MAFEIP's first report on outcome indicators, survey data indicated that additional coverage achieved through specific secondary indicators may be even lower than for common secondary indicators. For cognitive decline, functional status, frailty and adherence, it oscillated between 0% and 4% in all but Action Group A2. In this Action Group, only falls indicators sustained their relevance in terms of additional coverage of four responses to the survey (18%). Nutrition, which we handled as a special case in MAFEIP's first report on outcome indicators, only increased coverage of additional responses by 2%. Additional coverage was also low for instruments measuring and valuing mental health, in particular depression (1%), but for some of these instruments, mapping towards more general measures of HRQoL, such as the EQ-5D instrument, may be an option [e.g. Dakin, 2013].

We considered the inclusion of secondary common and specific indicators in MAFEIP's first report on outcome indicators to provide an alternative should we not find enough support for primary indicators. However, this support was overwhelmingly in favour of primary indicators in the survey. This survey was answered by 36% of participants across the six Action Groups and may therefore be broadly indicative of the data on indicators that may actually become available in the future. We decided, however, not to drop secondary indicators, but to consider the trade-off between the resources required for adding additional indicators to the framework, and the incremental increase in coverage that could be achieved, and the implications for the methods to link individual indicators with EIP on AHA objectives. Hence, in this report our aim is to develop a generic model that can be adapted to different target populations, interventions and settings, and which may be fed with data related to the different (primary and secondary) outcome indicators mentioned in Table 1 above. Furthermore, we have developed a section on potential extensions to the general model which would allow the use of additional data from those secondary indicators which the general model could not accommodate. However, further work in this direction will depend on the feasibility of and resource availability for developing such model extensions for linking up these secondary indicators to the EIP on AHA objectives.

1.3.2 "Sustainability of health systems" Column

In MAFEIP's first and second reports on outcome indicators, we proposed collecting information on the change in quantities of resources used as the primary focus of analysis for assessing the impact of EIP on AHA on the sustainability of health and care systems. The most frequently mentioned outcome indicators for the sustainability of health and care systems are:

- **Health and care resource use**, e.g. number of visits to primary care; measures of hospital / emergency admission and length of stay.
- **Health and care cost / expenditures** (resource use valued in monetary units) e.g. drug cost, administrative cost.

The above is supported by the scientific evidence which shows two main types of outcomes: cost-related and utilisation-related, as reported in MAFEIP's first report on outcome indicators.

We think that commitments should be encouraged to report any relevant resource use information, regardless of whether resource use items are likely to be positively or negatively affected by the

[6] The booklets can be found in the library of the EIP on AHA marketplace: https://webgate.ec.europa.eu/eipaha/library/

intervention. This relates to both changes in resource use patterns due to the intervention and also changes in the likelihood of related future health and care utilisation. On the other hand, collecting information on health and care cost / expenditures only (i.e. resource use valued in monetary units) may limit the possibilities of transferring or extrapolating information from individual commitments to other settings, or to a health/care system level. The main reason is that relative prices may differ substantially between settings, so that an identical change in resource utilisation in two settings may translate into differing impact on respective health and care expenditure. In addition, unit costs of resources are also likely to change over time within the same setting, which may impact on respective health and care expenditure even if utilisation remains constant. Therefore, whilst resource use data should be collected on an intervention level, we propose that information on unit cost for resources should be collected via desk research unless this data can be provided by EIP on AHA participants.

1.4 Aim and objectives of this report

After having defined a shortlist of candidate indicators for the quantitative assessment of impact within the EIP on AHA policy initiative in MAFEIP's first and second reports on outcome indicators, the next step in this project is to conceptualise a quantitative model that could be suited to establishing a link between candidate indicators and the EIP on AHA objectives. In previous reports, in particular in MAFEIP's first report on outcome indicators, we have already briefly considered the feasibility of building quantitative models to link potential outcome indicators with EIP on AHA objectives. However, we only did so to make a decision on the inclusion or exclusion of indicators, as their usefulness for MAFEIP depends on whether they a) can capture impact on an intervention level and b) allow a quantitative assessment and, at least in principle, extrapolation towards the EIP on AHA objectives. Hence, the primary aim of this report is to develop further the ideas expressed in MAFEIP's first report on outcome indicators and to conceptualise a quantitative approach for the estimation of impact using the outcome indicators previously identified as potential candidates for MAFEIP. This includes both the development of an overall conceptual framework and the provision of further details regarding the modelling strategy to link individual outcome indicators with the EIP on AHA objectives. However, it should be noted that because the EIP on AHA policy initiative is in its early stages and consequently actual outcome data is scarce, this report does not aim to provide final models to quantify EIP on AHA outcomes, as this will only be possible once we know exactly what kind of data will become available on an intervention level. Neither do we aim to provide a comprehensive framework that covers all possible cases (interventions / commitments) within the EIP, as this would be unrealistic given the size of the EIP on AHA and the varying scope of participating commitments. Furthermore, though the ability to link up to the EIP on AHA objectives played a role when we selected candidate indicators, the technical feasibility of doing so within the scope of MAFEIP is still unclear for some secondary indicators. Even though this question is not within the scope of the current deliverable, we will address both the relevance of some secondary indicators (such as, for instance, 'falls', 'frailty', or 'cognitive decline') for adapting the general model developed here to a particular target population and intervention, and also potential model extensions for the inclusion of other secondary indicators (such as, for instance, 'physical activity', 'risk factors', or 'adherence to treatment') into the framework. Hence with this deliverable, we aim to provide a methodological framework that tries to address the major challenges within MAFEIP, i.e. the mentioned variety of interventions within commitments and Action Groups and the expectation of very scattered and incomplete outcome data. Both require a highly flexible modelling approach that can be adapted to a variety of commitments, interventions, disease areas and geographic domains, whilst still providing reliable and scientifically robust estimates. Obviously, our proposal will require further fine-tuning once actual data becomes available to model the outcomes of commitments participating in the EIP on AHA policy initiative.

1.5 Structure of this report

The rest of the report is structured as follows:

- Chapter 2 provides a rationale for our overall modelling approach, which builds upon the general method to calculate Healthy Life Years (HLY). We will then provide details and examples of the calculation of health expectancies in general and HLY in particular. We conclude Chapter 2 by highlighting some of the key characteristics of the general HLY-model. This model may require adjustment for MAFEIP, which aims to adequately estimate the health outcomes achieved within the EIP on AHA.
- Chapter 3 will pick up on this discussion by trying to address these challenges whilst extending the general approach of calculating HLY and transforming it into a model which is more suitable for the purposes of MAFEIP.
    - Section 3.1 provides the basis for proposing a Markov-Model with three generic health states ('baseline health', 'deteriorated health' and 'death') to link outcome indicators across populations, interventions, commitments and geographic domains to the EIP on AHA objectives.
    - Section 3.2 elaborates on issues related to the implementation of the proposed model, such as the incremental approach that we adopt for MAFEIP, the development of a web-based tool based on the proposed model to enable commitments to assess their respective impact on the EIP on AHA objectives related to quality of life and the sustainability of health and care systems, the optional inclusion of additional health states to further improve the flexibility of the proposed tool, as well as data needs and potential sources to populate the tool with the required data.
    - Section 3.3 describes how some secondary outcome indicators (such as information on 'falls', 'frailty', or 'cognitive decline') could be utilised to adapt the generic Markov Model to a particular target population and intervention. We also discuss potential model extensions for other secondary outcome indicators (such as 'physical activity', 'risk factors' or 'adherence to treatment').
    - Section 3.4 explains how to adapt further the proposed model in order to estimate the impact of EIP on AHA activities on the sustainability of health and care systems.
- Chapter 4 discusses remaining challenges for quantitative modelling within MAFEIP.
- Chapter 5 offers our conclusions.

2. Adapting the Healthy Life Years (HLY) model to build a quantitative model for the EIP on AHA objectives

In this chapter, we first present the rationale for our approach, which consists in developing a model based upon the general method for the calculation of the HLY-statistic, since HLY have previously been chosen as the unit of measurement for the EIP on AHA headline target. We then provide details on the calculation of health expectancies in general and HLY in particular, before we elaborate on some characteristics of the HLY-statistic that need special consideration for the purposes of MAFEIP. Attention will be given to the difference between Healthy Life Years (HLY) and Quality Adjusted Life Years (QALYs), as we deem it crucial to incorporate information on health related quality of life (HRQoL) in order to estimate the potential impact of the EIP on AHA on its Quality of Life objective.

2.1 Rationale for adapting the HLY model

In the MAFEIP first and second reports on outcome indicators we outlined the major difficulties with extrapolating health outcome within the EIP on AHA towards the HLY-statistic. In brief, the analysis done in these reports showed that the HLY indicator is not actually used by any commitment of the EIP on AHA. Further, in the "MAFEIP First report on outcome Indicators" we discussed some key characteristics of the HLY-statistic that reduce its capacity to capture potential health gains from the activities delivered within the EIP on AHA. These characteristics may not be regarded as a disadvantage of the HLY-statistic per se, but rather as a consequence of the fact that HLY have been developed with a different purpose in mind than estimating incremental health gain from a particular intervention. In the "MAFEIP First report on outcome Indicators" we have therefore proposed to extrapolate from outcome indicators in the QoL block of the monitoring framework to Quality Adjusted Life Years, and to add gains in quality adjusted life expectancy across interventions and commitments targeted at different populations. We have further stated that, for the final link between the QALY-statistic and the headline target expressed in HLY, more exploratory work will be required.

As a consequence, we have decided - in a first step - to systematically review the scientific literature to help bridging between those two measures of health expectancy, Healthy Life Years and Quality Adjusted Life Years. The review, conducted between March 18th and March 24th 2014 across SCOPUS, Pubmed, Cochrane Library and Health Economics Evaluation database (HEED), aimed to identify papers that have previously discussed or attempted to provide a link between the HLY and the QALY metric. We have further extended our search to include papers looking into potential links between Disability Adjusted Life Years (DALY) and HLY, hoping this could result in further insights about potential links between indicators that could be useful for our purposes. Unfortunately, we could not identify suitable models in the state-of-the-art scientific literature that would provide such linkage. This may be due to the above mentioned fact that HLYs and QALYs have been developed with a different purpose in mind. Whilst QALYs are useful for estimating health gain from particular healthcare interventions [e.g. Weinstein et al., 2009], which makes them particularly suitable for the purposes of MAFEIP, the primary aim of the HLY statistic is to provide information on the current health status (or rather health expectancy) of a particular target population [e.g. Robine et al., 1999; Jagger et al., 2008]. Appendix 1 provides details on the search strategy used within each database, the search results, and the number and references of papers reviewed in full-text.

Whilst not providing an actual model for linkage between the QALY and the HLY statistic, the literature review did result in a number of useful insights for understanding the construction of HLY. Hence, we have used this information to look into the details of the method used for building the HLY indicator in order to find out whether this could give us some clues as to how we may be able to bridge the link between Quality Adjusted Life Years on the one hand and Healthy Life Years on the other.

In brief, both HLYs and QALYs are measures of health expectancy which combine information on the length of life with information on the health state in which this life is being lived (e.g. Weinstein et al.,

2009; Jagger et al., 2006). At their extreme values (i.e. 'full health' and 'death'), both measures may be regarded as equivalent, that is, one year in full health should be equivalent to one HLY or one QALY. Respectively, 'death' can be expressed as 'zero' HLYs or 'zero' QALYs. However, the key difference between both measures lies in the continuum between 'full health' and 'death'. Whilst the HLY-statistic only captures one state of 'disability', which is defined as being limited in usual activities for at least six month (Jagger et al., 2006, Eurostat: EU-SILC, 2014), QALYs allow discriminating between health states on an interval scale between 0 ('death') and 1 ('full health') (e.g. Weinstein et al., 2009). Further the use of QALYs is advocated on the basis that, from a public policy perspective, their value is the same, no matter who gets it (e.g. Weinstein, 1988). These properties allow us to argue that an incremental QALY gain for many should be equivalent to a large gain in QALYs for a few, i.e. 100 * 0.01 QALY = 1 QALY = one year in full health = 1HLY. This makes the QALY statistic particularly useful for MAFEIP as it has the potential to overcome one major drawback of the HLY-metric, which is its insensitivity to capture health gains which neither result in prolonged life nor fully restore the health of an individual (i.e. move the individual from an 'disabled' state to the 'healthy' state). In other words, if patients gain from the interventions delivered within MAFEIP, but not sufficiently for their health to be restored (as it is the case by definition for chronically sick patients for instance) the HLY-statistic would not, in contrast to the QALY, be able to capture this. On the other hand, if incremental QALY gains add up to a full QALY, and a year in full health is the same irrespective of our choice of outcome measure, the distinction between the QALY statistic and HLY becomes, at least in theory, redundant.

In practice, however, there are differences between the two which have to be considered further. It is obvious that we cannot estimate HLY in MAFEIP: on the one hand this would not allow capturing the above mentioned incremental health gains that do not result in full health of patients treated within the EIP on AHA, on the other we simply lack the data required to estimate HLY in the way EUROSTAT does as none of the commitments participating in the EIP on AHA reported to estimate HLY. Hence, any aggregated measure of health outcome will inevitably differ from the HLY-statistic as deployed by EUROSTAT. The best we can do, under these circumstances, is to take the HLY-model as a starting point and develop our own model by adapting the HLY-statistic to the specific needs of the EIP on AHA monitoring and assessment framework, most importantly by including HRQoL data to allow estimating health gains on an interval scale. This approach strikes a balance between the need to obtain data on health gain that can be easily understood, communicated and connected to the EIP on AHA headline target (i.e. the 2HLY gain and Quality of Life improvement), and the required adjustments to the HLY model that we deem necessary to allow for efficient monitoring and assessment of health gain within the MAFEIP project. This also means, however, that there are constraints with regard to the comparability between the metric to express overall health impact of the EIP on AHA as developed in this proposal, and the HLY-statistic as deployed by Eurostat to estimate the health expectancy of European citizens [Eurostat: glossary, 2014]. We will discuss these differences in detail throughout the remainder of this document.

The following sections will explain in more detail how the HLY-statistic as deployed by Eurostat is constructed and which features may require adaptation for our purposes.

2.2 Constructing Healthy life Years (HLY) as a measure of health expectancy

Health expectancies are "*the combination of a life expectancy figure with a concept of health which makes it possible to calculate the number of years lived in different health states*." [Robine et al, 1999] In other words, health expectancies generally comprise the length of time spent in different health states, no matter how health states may be defined [Jagger et al, 2006]. As Robine et al. (1999) further specify, there are "*as many possible health expectancies as there are concepts of health*." One particular type of health expectancy is the Healthy Life Year statistic (HLY) as currently deployed by Eurostat [Eurostat: glossary, 2014], which was also chosen as indicator for the headline target of the EIP on AHA.

The HLY indicator deployed by Eurostat belongs to the European Core Health Indicators (ECHI) and is defined as the number of years that a person is expected to continue to live in a healthy condition [Eurostat: glossary, 2014]. The indicator is calculated separately for men and women and for three ages (at birth, at the age of 50 and at the age of 65 respectively) [Eurostat: glossary, 2014]. Mortality statistics (*length of life/life expectancy*) are being combined with data on self-reported activity limitation/disability (*health state*). Whilst mortality data is obtainable from the Eurostat demographic database [Eurostat: population data, 2014], self-reported disability data stems from the EU Statistics on Income and Living Conditions (EU-SILC) [Eurostat: EU-SILC, 2014], which contains a module on health. Respondents are being asked: "for at least the past 6 months, to what extent have you been limited because of a health problem in activities people usually do? Would you say you have been 1. severely limited / 2. limited but not severely, or / 3. not limited at all?" [Eurostat: EU-SILC, 2014]. The HLY statistic is then calculated following the 'Sullivan method' [Eurostat: glossary, 2014].

The Sullivan method is described in great detail in work published by the European Health Expectancy Monitoring Unit (EHEMU) [Jagger et al., 2006]. In brief, health expectancies calculated by the Sullivan method take into account the years that a person of a particular age can expect to remain living in a particular health state (no matter how this state may be defined) [Jagger et al., 2006]. To calculate HLY according to the Sullivan method, information on the age-specific prevalence of particular health states among the population is required, which in our case is provided by the EU-SILC data [Eurostat: EU-SILC, 2014]. For the HLY statistic, the two 'disability' states (severely and moderately limited) are merged into one state, leaving us with three potential health states to consider: 'healthy', 'disabled' and 'death' [Jagger et al., 2006]. For estimating the transition into the 'death' state we require mortality data, which is, as mentioned above, available from Eurostat's demographic database [Eurostat: population data, 2014]. Age cohorts are often expressed in five or ten year intervals, as proportions of different health states may be too small in smaller age intervals [Jagger et al., 2006].

The Sullivan method is based on a life table approach as illustrated in Table 2 below:

Table 2: Life table approach for constructing Healthy Life Years (HLY)

1	**Age**	X
2	**Mid- year population size**	P_x
3	**No of deaths**	D_x
4	**Central death rate**	$m_x=D_x/ P_x$
5	**Conditional probability of death**	$q_x = \frac{m_x}{1 + (1 - a_x) * m_x}$
6	**Numbers surviving to age X**	$l_{x+1}= l_x(1-q_x)$
7	**Person years lived at age x**	$L_x=0.5*(l_x+l_{x+1})$
8	**Total number of years lived from age x**	$T_x =\sum L_x$
9	**Total life expectancy**	$e_x=T_x/l_x,$
10	**Proportion with disability**	π_x
11	**Person years lived without disability at age x**	$(1-\pi_x)L_x$
12	**Total years lived without disability from age x**	$\sum (1-\pi_x)L_x$
13	**Disability free life expectancy**	$DFLEx = \frac{\sum (1 - \pi_x)L_x}{l_x}$
14	**Proportion of life spent disability free**	$\%DF = DFLE_x / e_x$

Adapted from Jagger et al (2006)

In order to populate Table 2 with data and to construct HLY, we require information on the three rows shaded in grey, i.e. the population size (P_x, row 2), the number of deaths in each age cohort (D_x, row 3), and the proportion with a disability (π_x, row 10) in each age cohort [Jagger et al., 2006]. With this data, we can calculate the central death rate (m_x, row 4), which is simply $m_x=D_x/ P_x$. The conditional probability of death in the age cohort is then calculated as $q_x = \frac{m_x}{1+(1-a_x)*m_x}$, where a_x is a half-term correction as people who die in period x usually have lived throughout part of that period (therefore, a_x is often set constant at 0.5 [Jagger et al., 2006]. For l_x, we may, for instance, assume a starting population of 100.000 in the first year of the life table. For successive years, we calculate the number of people surviving to age x as: $l_{x+1}= l_x(1-q_x)$. The number of person years lived at age x is then the average of the respective age interval and the subsequent interval, hence $L_x=0.5*(l_x+l_{x+1})$ (apart from the final age interval, and the first year of life). For the final year of the life table, L_x is given by l_x/m_x. T_x is the total number of years lived from a particular age, and simply the sum of L_x from that particular year up to the end of the life table. Finally, total life expectancy at each age is given by $e_x=T_x/l_x$, and this is displayed in the red shaded line 9 of table 2.

In order to move from life expectancy to disability-free life expectancy, we now need to multiply the person years lived at a particular age (L_x) with the probability of living without disability in that age cohort ($1- \pi_x$). The total number of years lived without disability from age x (row 12) is then simply the sum of ($1- \pi_x$)*L_x, from a particular age up to the end of the life table, equivalent to calculating the total number of years lived in a particular age T_x in row 8. Finally, by dividing the number of years lived without disability (row 12) by the number of people surviving up to a particular age l_x (row 6), we obtain the disability-free life expectancy for that age cohort. Dividing disability-free life expectancy by total life expectancy results in the proportion of remaining life spent disability free (row 14). A more thorough discussion and examples of how to calculate HLY are obtainable from [Jagger et al., 2006].

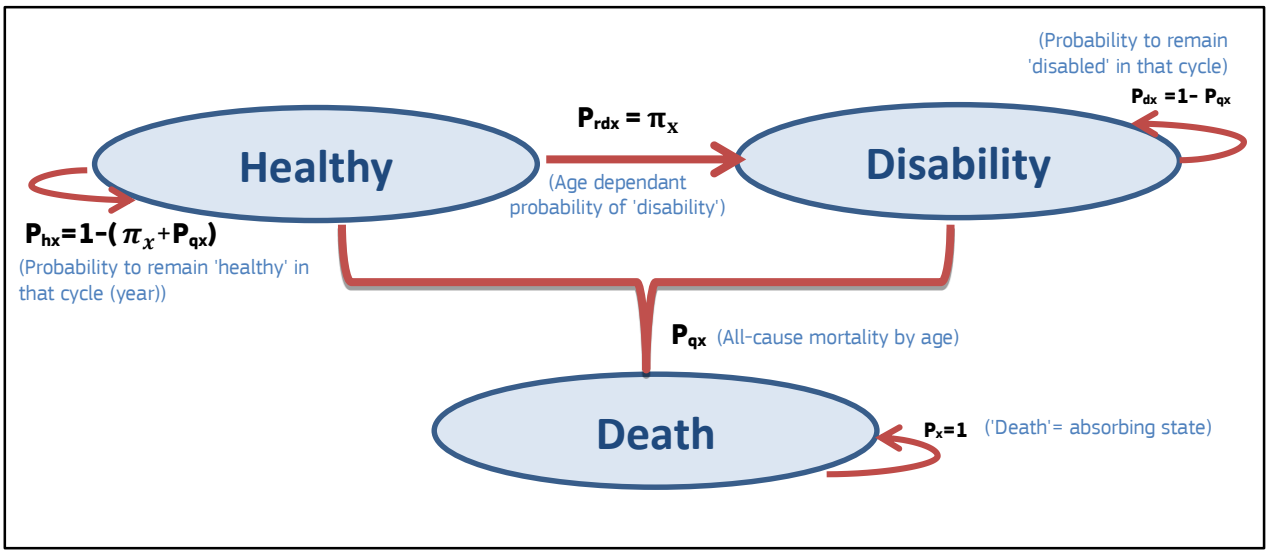
2.3 Problems with capturing potential health gains within the EIP on AHA through the standard HLY-model

What we can learn from the above section is that HLY are calculated using a cohort model implemented as a life table that rests on transition probabilities between three health states ('healthy', 'disabled' and 'death'). To populate the general HLY-model, three types of data are required:

- population size,
- the death rates in each age cohort, and
- the probability of becoming disabled in each age cohort.

Figure 2 provides a schematic representation of this model. The bubbles represent health states and the arrows the transition probabilities between those states. The schematic presentation displays one cycle of the model (i.e. one year of the life table displayed in Table 2 above), and the cohort (e.g. 100,000 patients) is assumed to progress through the model depending on the transition probabilities for a specified number of cycles (years). The longer the model runs (i.e. the more cycles we allow for), the higher the percentage of individuals that is progressing into the 'death' state, which is also referred to as an 'absorbing state' (as the probability (P) of remaining in the 'death' state is 1). Initially, the percentage of individuals of the cohort progressing into the 'disability' state is also increasing; however, as individuals ultimately move on to the 'death' state, this percentage gradually reduces from some point onwards over the remaining model-lifetime.

Figure 2: Schematic presentation of the HLY-model according to the Sullivan-method



From Figure 2, we can see that the HLY model only considers the all-cause mortality of a patient cohort stratified by age; and it does not decompose the probability of moving to the 'disabled' state from the probability of moving back to the 'healthy' state (i.e. the chance that an individual may fully recover from a particular condition). Related to that is the problem that the HLY model cannot be adapted to populations which are not fully healthy at baseline and whose health condition may deteriorate over time (e.g. patients with chronic diseases). Finally, the HLY model does not assign different weights to different health states according to the perceived quality of life in this particular state. The remainder of this section will discuss, in more detail, why these features may reduce the flexibility of the model and also potentially reduce its capacity to capture health impact of the activities carried out within the EIP on AHA. Note, however, that the above remarks do not aim to criticize HLY which were originally developed for an entirely different purpose (which is the provision of population based statistics on health expectancy). We merely highlight those features of the HLY-model as they may require some adjustment for the purposes of MAFEIP to allow for the monitoring of health outcomes within the EIP on AHA.

In order to understand better the limitations we face, let us consider the actions carried out within the EIP on AHA and their potential to improve the health of participating individuals. In general, potential health gains can be of the following types (note that interventions may also have a simultaneous effect on any of these categories):

I. An intervention could target healthy individuals to delay their progression into the 'disabled' state of the HLY model, i.e. their probability (P) of declining health **(P-decline)**. This could be the case, for instance, if a commitment focusses on healthy elderly people to increase their general physical activity levels. As a consequence, the probability of adverse future health conditions such as, for instance, cardiovascular disease (CVD), decreases. This, in turn would lead to prolonged healthy life, and therefore contribute towards the EIP on AHA health target by slowing down the progression into the 'disabled' state.

II. On the other hand, an intervention may target patients who already suffer from a particular condition in order to fully restore their health **(P-improvement)**. In other words, some patients may move from the 'disabled' state back to a 'healthy' state

III. Health gain may also occur albeit without full health recovery. This means that patients move from the 'disabled' state to another health state which is characterised by higher **health related quality of life (HRQoL),** but which is not equal to the 'healthy' state of the HLY-model. Likewise, for patients who are not healthy at baseline, the objective could be to prevent

a worsening of the current health status so as to avoid moving to a health state which is characterised by lower HRQoL (and potentially higher health and care resource use). This may be the major source of health gain for any commitment focussing on patients with chronic diseases within the EIP on AHA as their probability to move from / to fully healthy life is, by definition, zero. The HLY-model is not able to adequately capture such situations without further adaptation.

IV. Finally, an intervention may also change the probability of death (**P-death**) of either healthy individuals (e.g. through primary prevention) or patients who already suffer from a particular health condition. For instance, primary prevention of CVD could be targeted at healthy individuals so as to lower their probability of a sudden fatal CVD-event, whilst secondary CVD prevention, which targets patients who already have a history of cardiovascular disease, could lower the disease specific probability of death for this cohort.

Hence, in order to adapt the HLY model for MAFEIP, we face a number of issues: first and foremost, a suitable model for MAFEIP needs to be adaptable to a large number of commitments within the six thematic Action Groups of the EIP on AHA, which focus on a variety of objectives, implement different interventions and target different cohorts of individuals with different demographic or disease characteristics. Hence, the HLY-model with its three rather restrictive health states ('healthy', 'disabled', 'death') does not provide the flexibility that would be required to allow adaption to various contexts. For instance, the HLY model is predicated on the assumption that recovery from the disability state should always result in full health. Hence, any intervention delivered within the EIP on AHA that does improve health without fully restoring it, or that focusses on patients with a particular condition (such as chronically sick) and aims to reduce the transition into a worse health state, would not be captured by the HLY-model introduced above. Therefore, a more generic definition of health states is required. To achieve the flexibility needed for MAFEIP, we will introduce in the next chapter three generic but mutually exclusive health states for the model: 'baseline health', 'deteriorated health', and 'death'. Adapting such health states to a particular patient cohort (e.g. patients at different stages of frailty as measured by the Groningen Frailty Index, GFI [Steverink et al., 2001]) and weighting them with respective estimates of HRQoL (measured on an interval scale) would then allow adding up health gains across a diverse set of individuals, interventions, and contexts.

Secondly, as the standard model used by Eurostat to calculate HLY according to the Sullivan method does not decompose explicitly the probability of disability (π_x) from the probability of recovery [Jagger et al, 2006], it is not possible to explicitly model the fact that patients may (in part or even fully) recover from an impaired health condition as a consequence of actions being taken within the EIP on AHA. Of course, one may argue that the parameter π_x in the HLY-model (π_x = age dependant probability of disability) includes both the probability of disability and the probability of recovery. However, without decomposing π_x, it may not be possible to adequately consider health gain of individuals who may restore their health as a direct consequence of an EIP on AHA intervention.

Thirdly, as the HLY model used by Eurostat does not distinguish between the mortality rates of individuals in different health states [Jagger et al., 2006], it is not possible to quantify the effect of interventions that are able to prolong life by reducing the mortality associated with a particular disease. This holds both for healthy patients whose probability of a sudden fatal health event may change e.g. through primary prevention measures, as well as for sick patients whose disease specific mortality may change due to a particular intervention. Hence, for the purposes of MAFEIP, we require a model that decomposes the mortality of individuals in different health states so that any change due to an intervention may translate directly into health gain attributable to the EIP on AHA.

Finally, and this is probably the most fundamental issue with the HLY-statistic for the purposes of MAFEIP, we found that none of the commitments participating in the EIP on AHA collect the information required to construct HLY. This could be a consequence of the above mentioned limitations of the HLY-statistic for the purposes of monitoring health outcome of individual health interventions.

The following chapter will pick up from this discussion and show that some adjustments to the HLY-model may help address the above mentioned issues, making the model better suited for estimating health gains from the activities delivered within the EIP on AHA. It will further show that the resulting model can be implemented as a general three-state Markov-cohort model, which is based on the two-primary outcome indicators (HRQoL and mortality) previously identified for the purposes of MAFEIP (see MAFEIP first and second reports on outcome indicators).

3. MAFEIP quantitative modelling proposal

In this chapter, we will expand on the HLY-model and propose a general health-state transition model for quantifying health outcomes within MAFEIP. We have identified Decision Analytic Modelling (DAM), an approach that is commonly used for the economic evaluation of healthcare technologies [e.g. O'Brien 1996; Drummond et al., 2005, Briggs et al., 2006], as a promising strategy to implement models that could be suited for the EIP on AHA monitoring and assessment purposes. More precisely, we will present and discuss a general three-state Markov cohort model for the quantitative linkage of previously identified outcome indicators with the EIP on AHA objectives (see Section 3.1). The model is built around the two primary outcome indicators (HRQoL and mortality), but those secondary outcome indicators that could be used to describe mutually exclusive health states for the model (such as different stages of frailty, cognitive or functional decline) are also relevant and may be used to adapt the model to a particular target population. Section 3.2 will then elaborate on some issues related to the implementation of the proposed model, such as the incremental approach that we adopt for MAFEIP, the development of a tool based on the proposed model to enable commitments to assess their respective impact on quality of life and health and care systems' sustainability outcomes, as well data needs and potential sources to populate the tool with the required data. Whilst some secondary outcome indicators may be useful and required to define health states when adapting the tool to a particular target population, other secondary indicators which were discussed in previous reports (such as physical activity, risk factors or adherence rates) may require extensions to the generic Markov model. We will briefly discuss potential model extensions in Section 3.3 in order to show how such secondary indicators could, in principle, be incorporated into the framework. We will conclude Chapter 3 by showing how the proposed framework could be adapted further to estimate the impact of EIP on AHA commitments on the sustainability of health and care systems in Section 3.4.

3.1 Extending the standard HLY-model for the purposes of MAFEIP

As discussed in Section 2.3, we have identified a number of issues associated with the standard HLY-model when trying to capture and aggregate health outcomes within MAFEIP:

- The model cannot explicitly capture changes in the probability of moving forth and back between its two 'alive' states ('healthy' and 'disabled')
- The model does not distinguish between baseline mortality and the excess mortality related to a certain health condition.
- The model fails to capture health improvements which neither result in full recovery nor apply to individuals who were not 'healthy' at baseline.
- The model does not possess the flexibility to be adapted to a diverse set of commitments within which 'alive'-states may be defined differently, and neither does it allow aggregating respective health impacts across different interventions, populations, or contexts for which the various alive states differ.
- None of the EIP on AHA commitments collect the information on self-reported disability, which is required to estimate HLY.

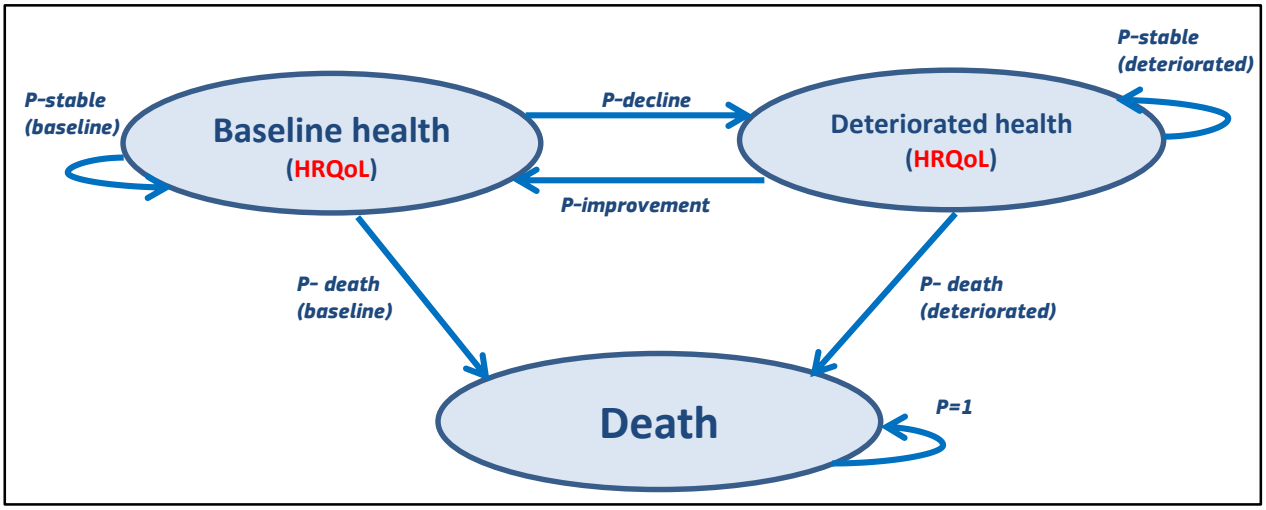
In order to overcome these challenges, we propose the following adjustments to the model:

1. Defining three generic and mutually exclusive health states: two 'alive'-states ('baseline health' and 'deteriorated health') and the absorbing 'death'-state respectively.
2. Explicitly modelling the probability to move forth and back between the 'alive' states of the model.
3. Decomposing mortality to allow for explicit modelling of changes in the probability of 'death' for individuals in different 'alive'-states as a consequence of an EIP on AHA activity.
4. Using HRQoL-data as reported by a large number of commitments across EIP on AHA Action Groups to weight different 'alive' states of the model. This will allow us to:

a. account for incremental gains in HRQoL within different 'alive'-states, and
b. aggregate health impact across different commitments that define respective 'alive'-states differently.

The resulting model is displayed in Figure 3 below. The baseline health state represents the HRQoL and life expectancy of the population targeted by the intervention. The deteriorated health state represents a condition where one or more morbidities have negatively affected an individual's HRQoL and life expectancy compared to the baseline health. The specific definition of this state will be determined by the nature and purpose of the intervention under assessment, as it represents the state of deteriorated health that the intervention aims to prevent or cure. Each state will have a HRQoL utility weight attached to it. Hence, the model is built around the two primary outcome indicators (HRQoL and mortality), however, it also allows adapting the "alive" states to different target cohorts, interventions and contexts by using data from secondary outcome indicators to define mutually exclusive health states for the model (such as, for instance, different levels of frailty as measured by the Groningen Frailty Index, GFI [Steverink et al., 2001]).

Figure 3: Generic Markov-model to enable quantification of EIP on AHA health impact



The simulated population will transfer between the Markov states based on four transition probabilities: the incidence of the deteriorated health condition, the rate of recovery from that back to baseline health, baseline mortality in the target population, and excess mortality in the population with deteriorated health. The resulting model is more flexible and may allow a more comprehensive assessment of the possible health gains that may be achieved through interventions delivered within the EIP on AHA. By weighting different 'alive'-states using HRQoL data, we draw upon the information likely to become available across a large number of commitments and Action Groups within MAFEIP (as previously assessed in the MAFEIP first and second reports on outcome indicators), and we adequately reflect the fact that different commitments (and often different interventions within commitments) may relate to entirely different health conditions.

In order to reflect disease progression more adequately within a health state transition model, however, one could also define more health states, and different HRQoL-weights could then be assigned to each state [e.g Briggs & Sculpher, 1998; Briggs et al., 2006]. The optional inclusion of additional health states (beyond the three states mentioned above) will therefore be explicitly considered when developing a web-based 'monitoring tool' for MAFEIP as this would further enhance the flexibility of the tool to be adapted to different interventions, target populations and / or care contexts. However, our main purpose in this section is to introduce a conceptual framework for an extrapolation model for MAFEIP which is (relatively) simple and flexible. The more complex the basic model, the more context-specific it may become, and the more resource consuming it may get to adapt

it to particular commitments [e.g. Haycox et al., 1998, Sculpher et al., 2004]. The economic evaluation literature offers a number of examples of generic models which have been adapted ex-post to different settings [e.g. Drummond, 1994; Haycox et al., 1998]. Our strategy within MAFEIP is similar, though the number of different settings to be considered is unparalleled. Further, whilst we believe that the optional inclusion of additional health states will enhance the flexibility of the tool to be adapted to different interventions, populations and contexts, this also means, however, that more parameters will have to be populated with data. This would only be feasible if such data is likely to become available for the respective intervention / commitment, as populating such models with secondary data from alternative sources would significantly increase the use of analytic resources (for more details on this issue see also Section 3.2). Data availability is likely to be a limiting factor within MAFEIP which we ought to take into account when adapting a model to a particular context. Thirdly, the model proposed for MAFEIP is originally based on the HLY-calculation method as described in Chapter 2, and all changes to the original HLY model should be justified and based on prior reasoning. Finally, we consider it crucial that the model can be communicated easily and clearly to the various stakeholders within and outside the EIP on AHA. However, to warrant capturing potential health gains within the different commitments of the EIP on AHA, some level of complexity is unavoidable especially if the model is to provide scientifically robust and unbiased evidence that is useful for policy making. At the same time the model should remain as straightforward as possible to allow non-experts to understand the methods that led to the evidence generated within MAFEIP. We believe that the general three-state transition model that builds on the HLY-model balances out this potential trade-off, but the optional inclusion of additional health states when adapting the model to a particular context could further enhance the flexibility of the monitoring framework and should therefore be considered when building a web-based monitoring tool for MAFEIP (see Chapter 3.2).

With respect to HRQoL, note that the analysis of survey data in the "MAFEIP Second report on outcome indicators" showed that commitments may use a variety of instruments to elicit information on patients HRQoL, most prominently the EQ-5D [http://www.euroqol.org/] and the SF-36 tool [Ware et al, 2000], but also other methods such as the SF-12, [Ware et al, 1996], the SF-6D [Brazier et al ,2002], the Nottingham Health profile [Hunt, 1981] or the 15D instrument [Sintonen, 2001], amongst others. As a consequence, HRQoL data may not be comparable across interventions if different methods have been used. This variability in methods constitutes an obstacle for the aggregation of health gains across the groups of interventions for which different instruments were applied. However, the scientific literature offers methods for 'mapping', or 'cross-walking' between different elicitation methods [e.g. Dakin, 2013]. Therefore, in order to increase the generalizability of our proposed framework to different contexts, and to be able to aggregate health gains across groups of interventions, we propose either calculating or adapting existing algorithms that are published in the scientific literature to 'convert' HRQoL-values measured with one instrument (e.g. the SF-36) into equivalent values of another instrument (such as the EQ-5D). There are both theoretical and practical reasons why mapping from other instruments towards the EQ-5D may be indicated [e.g. Brazier et al., 2010]. First, the EQ-5D was most commonly mentioned across commitments in the survey on outcome measures reported in the "MAFEIP Second report on outcome indicators". Secondly, the EQ-5D instrument is a generic, preference based measure, with value sets and translations for most EU-countries [e.g. Brooks et al., 2003], and it is also the method of choice of a number of decision bodies, most prominently the National Institute for Health and Care Excellence (NICE) in England and Wales [NICE, 2009]. In addition, it has been explicitly mentioned in national pharmacoeconomic guidelines, submission guidelines for pharmaceuticals or published pharmacoeconomic recommendations for the Baltic countries (Latvia, Lithuania, Estonia), Belgium, Croatia, Hungary, Ireland, Poland and Sweden, and it is also compatible with recommendations available for countries such as the Netherlands, France or Spain [ISPOR, 2014: Eldessouki et al., 2012]. This may also be the reason why algorithms for mapping from other instruments towards the EQ-5D may be the most prevalent in the related scientific literature [Longworth & Rowen, 2013; Dakin, 2013]. The discussion in Chapter 4 will further elaborate on the method of mapping and its potential use within MAFEIP.

Having defined a general model structure for our purposes, the question is how to use such a model to quantify potential health gain from a particular intervention carried out within the EIP on AHA in practice. This will be discussed in depth in the next section.

3.2 Implementing the proposed modelling framework

In this section, we will build upon the general model structure developed above and show how this can be used to assess incremental health gain within the EIP on AHA using Decision Analytic Modelling (DAM) techniques. When applying DAM to the health field, the overriding question is *essentially* what to do based on all the information which we *currently* have about a particular problem [e.g. Drummond et al., 2005]. The basic idea is that a decision cannot be avoided even if the information to support this decision is scarce, meaning that such decisions need to be taken *"under conditions of uncertainty*" [Drummond et al., 2005]. The basic strategy is to "*synthesise all available information from multiple sources and to apply mathematical techniques to assess the impact (costs and outcomes) of health interventions*" [O'Brien, 1996]. The essential data inputs are probabilities for clinical events, and the impact of such events on a) costs (resources valued in monetary units) and b) values or utilities for health outcomes [Obrien, 1996]. The appeal of this approach in general, and particularly for the purposes of MAFEIP, is the fact that DAM "*pulls together the many needed pieces of information from multiple sources and then stitches them together into a (hopefully) cohesive whole*" [O'Brien, 1996]. Hence, DAM is perfectly suited to dealing with a potential major challenge within MAFEIP, which is the availability of data across a diverse set of commitments / Action Groups as existing gaps in the data from commitments may be filled with data from other sources and the impact of parameter uncertainty on model results can be quantified [e.g. Briggs & Sculpher, 1998; Briggs et al., 2006].

It needs to be mentioned, however, that the European Commission does not intend to assess the incremental cost-effectiveness of an intervention carried out by an EIP on AHA commitment, nor to compare several interventions on their cost-effectiveness. Rather, the general aim of the MAFEIP project is to estimate the aggregated impact of the EIP on AHA on its overall quality of life and sustainability of health systems objectives. Nevertheless, a web-tool implementing the Markov model introduced above could be used by individual stakeholders to estimate the Incremental Cost Effectiveness Ratio or the Incremental Net Monetary Benefit of one intervention compared to another [Stinnett & Mullahy, 1998].

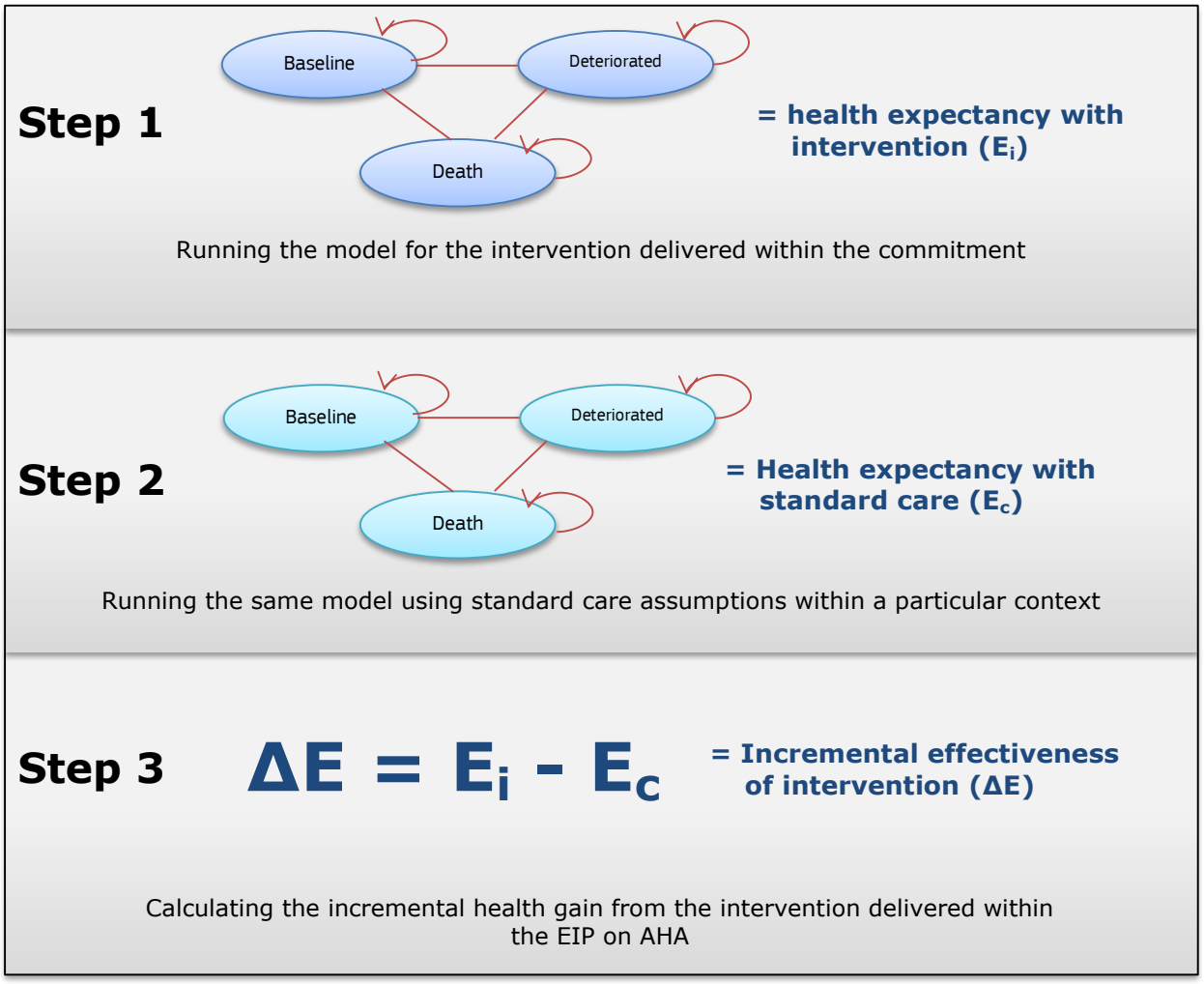
We would like to address now three key issues relating to the implementation of our proposal and the estimation of health impact within MAFEIP using DAM, namely:

i. How to calculate incremental health gain associated with activities in the EIP on AHA
ii. How to deal with contextual variation within and between commitments
iii. How to populate models with relevant data

To answer the first question, consider Figure 4 below. In order to estimate the health gain that results from an intervention delivered within the EIP on AHA, it is not sufficient to run the model developed above only once with parameter estimates for the respective intervention under assessment. This would result in an estimate of total health expectancy, which includes the health gain associated with the intervention, but without making it explicit [e.g. Drummond et al., 2005]. In other words, in order to find out how much the EIP on AHA activities *add* to the health expectancy of European citizens, we also need to know the health expectancy that would have been achievable without having the EIP on AHA in place [e.g. Drummond et al., 2005]. This is what has been referred to in previous communications as "*defining the baseline*" (henceforth we will only use the term "standard care scenario"). Only if we appropriately quantify health expectancy without the EIP on AHA, can we estimate the potential health gains from the interventions delivered within commitments and Action Groups. To do so, we need to subtract the health expectancy that would result from standard care in a particular context from the health expectancy with an EIP on AHA intervention in place [e.g. Drummond et al., 2005].

Hence, to estimate the "*incremental health gain*" from an intervention delivered within an EIP on AHA commitment the model needs to run twice, and this incremental health gain is then calculated as shown in Figure 4 as $\Delta E = E_i - E_c$ [e.g. Drummond et al., 2005] (where $E_i$is the health expectancy with intervention, and E_c the respective health expectancy with standard care). The total health impact across all EIP on AHA commitments is then $\Delta E_{EIP} = \sum (E_i - E_c)$.

Figure 4: Incremental approach to estimate EIP on AHA health impact



Having clarified that the EIP on AHA impact should be estimated as the "*incremental health gain*" from the activities carried out within commitments and Action Groups, the next question that inevitably follows is how to define, for each intervention respectively, an appropriate *"standard care scenario*". The standard care scenario should generally reflect routine practice, hence, reporting health outcomes of the most common care strategy that would be replaced by the intervention in question. A similar definition of the comparator has been adopted in many national guidelines for the economic evaluation of healthcare technologies in developed countries [Eldessouki et al., 2012; ISPOR, 2014]. However, this would require assessing 'routine practice' for each intervention, subpopulation and geographic setting, as there may be considerable contextual variation [e.g. Sculpher et al;, 2004; Goeree et al., 2007; Boehler, 2013; Boehler & Lord, 2015]. Likewise, interventions differ vastly between individual commitments, and sometimes even within commitments, e.g. for different population subgroups or across different geographical sites within the same commitment. These sources of variation could constitute one of the most fundamental challenges within the MAFEIP project, as the

EIP on AHA gathers more than 500 commitments across six thematic Action Groups, often with more than one intervention per commitment, and spanning all 28 EU-member states (and more). Consequently, the monitoring and assessment of EIP on AHA outcomes would inevitably lead to a huge amount of scenarios to consider, each characterised by a unique set of parameter estimates for the model. Even if alterations in structural model assumptions are being ruled out, diversity within and between commitments, care settings, and geographic locations would still have to be reflected by respective variation in parameter estimates. Having to deal with variation on such a scale may ultimately cast into doubt the practical feasibility of monitoring outcomes within the EIP on AHA in the above described manner.

For this reason, instead of attempting to build and populate models for each intervention centrally by IPTS, we propose to implement the monitoring framework for the EIP on AHA in such a way that participating commitments can assess themselves their respective impact on health and health and care systems' sustainability outcomes. This could be achieved through a user-friendly software tool implemented in a widely accessible software environment, which allows remote monitoring of EIP on AHA outcomes directly by the commitment, with the support and guidance of IPTS.

Such a tool should be based on the above described Markov-process and allow for a minimum of three health states ('baseline health', 'deteriorated health', 'death'), but potentially provide the option of incorporating additional health states if required. It should further consider transitions between health states in order to link primary outcome indicators (mortality and HRQoL) with the objectives of achieving a health gain and improved quality of life for European citizens. The tool could then be adapted to different interventions, populations and care settings by using secondary indicator data to define the two 'alive' states of the model. For instance, 'baseline health' could describe a particular patient cohort before the occurrence of a fall, whilst 'deteriorated health' describes the health and resource outcomes for the same population after a fall has happened. The transition between both states represents the likelihood of a fall in the target population. Likewise, different instruments which aim to assess the progression of a particular disease (such as, for instance, the Groningen Frailty Index [Steverink et al., 2001], or the Clinical COPD questionnaire [van der Molen et al., 2003]) could be used to define the baseline and deteriorated states of the model respectively for that disease, and HRQoL data from a generic instrument such as the EQ-5D could then be used to weight the respective health states of the model. Hence, the specific definition of the 'alive'-states will be determined by the nature and purpose of an intervention under assessment, and each state will have a HRQoL utility weight attached to it. A user-friendly interface should allow the use of the tool by individuals with little or no prior experience in decision analytic modelling; and last but not least it should also be possible to update, revise, or modify the tool in the future. To allow for remote data entry and retrieval of results, a web implementation may be most appropriate, especially in case of a wider roll-out.

Implementing the model described in this report by means of a web-based monitoring tool allowing remote data handling may address some of the major challenges within MAFEIP and yield important benefits, both for the monitoring and assessment of the EIP on AHA and for the stakeholders involved in this policy initiative. For monitoring purposes, we may benefit from the expertise of stakeholders who are the best experts in their own fields, and may best be able to gather, validate and select the evidence required to populate models (with the guidance of IPTS), while reducing the burden of data gathering for IPTS. Stakeholders would also benefit from being able to use a tool that helps them evaluate the impact of their intervention on patient's health and the sustainability of health and care systems in terms of respective changes in health and care expenditure. Whilst the European Commission does not intend to compare interventions on their cost-effectiveness, individual stakeholders could use the tool to estimate the Incremental Cost Effectiveness Ratio (ICER) or the Incremental Net Monetary Benefit (INMB) of one intervention compared to another. The use of the tool could be facilitated through dedicated workshops and through direct communication with commitments participating in the EIP on AHA.

There are examples in the literature, both for developing web-based Markov tools similar to what we propose for the purposes of MAFEIP, and for running workshops to facilitate the use of such instruments for participants with little or no background in economic evaluation and decision modelling methods. For instance, the Multidisciplinary Assessment of Technology Centre for Healthcare (MATCH) project [www.MATCH.ac.uk], a research collaboration between leading UK Universities (Brunel, Birmingham, Nottingham & Ulster) and a number of industrial and public sector partners, developed a web-based tool that allows users to model the transition of a population of patients through a series of health states over time, based on a Markov process [http://www.nottingham.ac.uk/match/research/tools/markovtoolmain.html]. The model provides flexibility in the sense that it allows choosing between 2 and 5 different health states and it can be run for a specified number of cycles and a specified size of population. Of particular interest for the web-based implementation strategy of the monitoring framework are the experiences with using tools of this kind in workshops for participants with little or no background in economic evaluation and decision modelling methods [Crowe et al., 2010: Craven et al., 2012]. The authors state that *'even though small companies and healthcare purchasers have little prior knowledge of health economics, the key issues can be rapidly absorbed to be applied in decision making*.' [Crowe et al., 2010].

However, there are also a number of issues which would require further research. For instance, the tool should be flexible enough so that it can be applied to different contexts, but gathering the required data for input parameters should remain at a manageable level for the stakeholders who may not have a particular background in health economic evaluation and/or decision analytic modelling. This may require, for instance, to distinguish between different sets of input parameters: those input parameters which characterise the interventions treatment effect may be more context specific and should generally be informed by data that stems from the commitment (unless such data is not available, in which case secondary data sources from the literature or comparable interventions implemented elsewhere may be considered - see further below in this section). For other parameters, it may be sufficient to consider variability only at national level. As a consequence, the tool could provide background data for those estimates that are generally required to populate the model and also assumed to vary only between countries (e.g. all-cause mortality stratified by age and gender; unit cost estimates for major resource use items stratified by EU-countries for the assessment of impact on the sustainability of health and care systems – see also section 3.4 below), and users could adapt the model to their geographic context by selecting pre-defined background data that would udpate automatically. For other, more context specific data, however, manual input will be required. Such input values could then be stored in a common database so that other users could benefit from an ever expanding pool of data to populate the model with relevant estimates. Also, the model should allow updating information as new evidence emerges. A major issue which inevitably follows relates to the loss of control over the quality of the data used to populate parameters, which in turn may have a negative impact on the validity and reliability of estimated outcomes. Hence, more research would be required to develop quality checks for data inputs, and also to develop appropriate guidance for potential users on the process of data entry.

To further specify the characteristics of such a tool, an iterative approach is recommended. Once a basic tool is developed and implemented in an adequate software environment, it should be tested in the field with a limited number of cases in order to:

- Assess its applicability to a number of potentially diverse interventions, target groups, and geographic domains, and provide suggestions to further improve its generalisability
- Evaluate the ability of the tool to estimate health and economic outcomes (and related uncertainty around point estimates) of the EIP on AHA.
- Incorporate and evaluate quality checks for data-inputs especially if the tool is to be used by stakeholders with limited background in economic evaluation in health.

- Test the tool and its interface with different stakeholders participating in the EIP on AHA and for a number of selected cases.

A critical appraisal of the initial tool should then form the basis for further refinement, before rolling it out on a wider scale. This appraisal should consider issues such as the time and resources required to populate the tool and the potential for improving remote data collection and robustness of results, e.g. through stakeholder workshops, training manuals, etc.

By proposing to implement the above described model as a tool for stakeholders participating in the EIP on AHA, we aim to address the issue of diversity arising from the wide scope of the EIP on AHA, with its six thematic Action Groups, comprising hundreds of commitments which focus on a vast variety of health and care interventions. It leaves us, however, with the third question raised at the beginning of this section, which is how to populate the tool with relevant data. It is likely that not all data required to populate the tool will be available from each commitment which may in principle be suitable for modelling, which is why DAM is particularly well suited, since its particular strength is to *'pull together'* the pieces of information needed from other sources [O'Brien, 1996]. This, however, brings us back to the importance of quality checks for data-inputs. As commitments should be able to use the tool to assess health and health and care systems' sustainability impacts of their respective interventions by entering their data remotely, the remainder of this section discusses the choice of input data especially for those parameters which commitments would generally have to populate themselves. Two questions arise:

- which sources to check for relevant data, and
- in which order should they be checked.

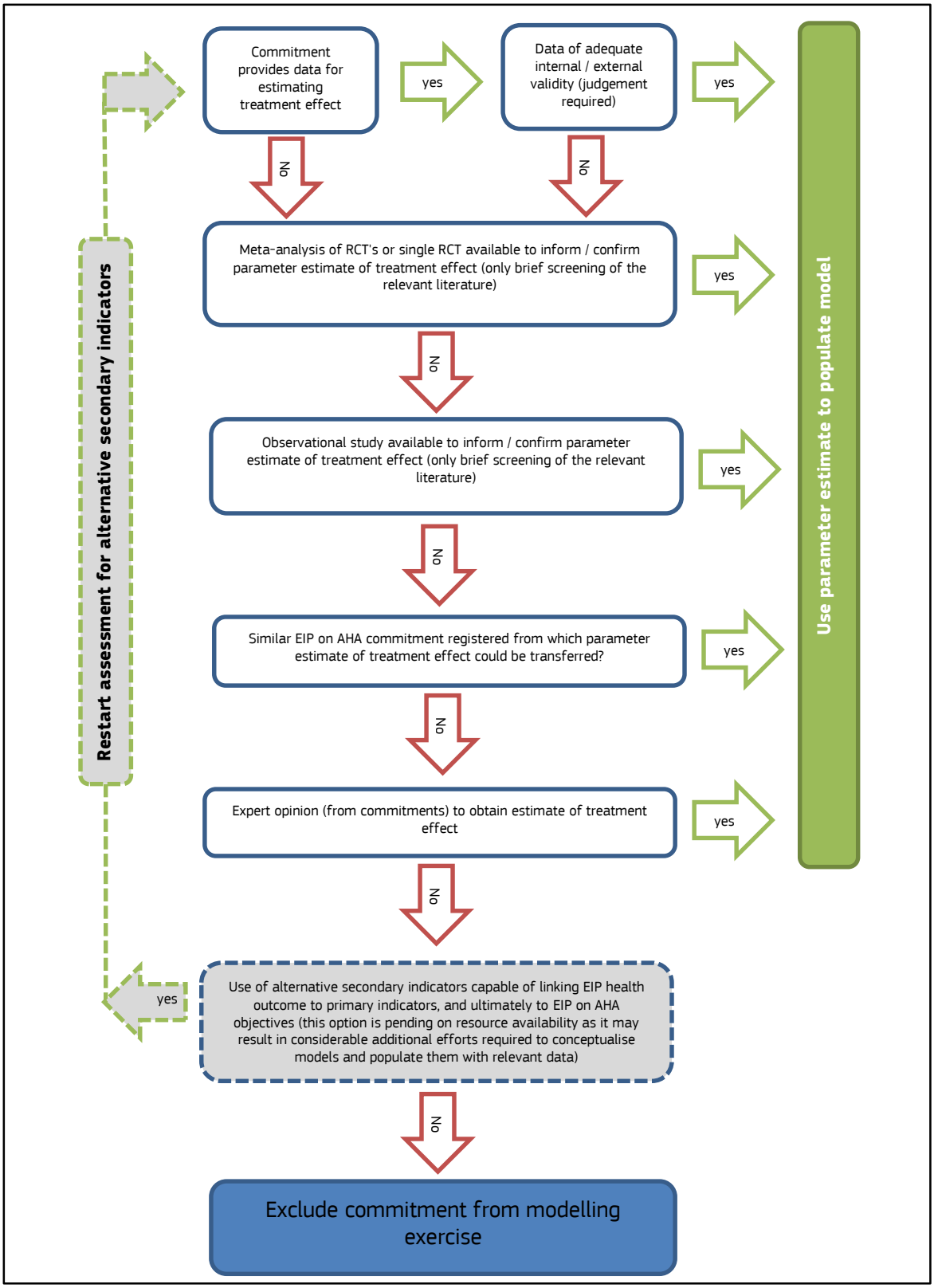
In order to populate the tool with relevant data, we should generally proceed from the most to the least adequate source of available evidence. Guidelines for the economic evaluation of healthcare technologies suggest particular sources of evidence for what is called treatment effects, i.e. the difference in health state (HRQoL) and the time spent in a health state (duration) between treatment and no treatment/placebo, or any other suitable comparator. For instance, NICE (2013) states that *'research methods and designs that are used to measure the treatment effect can broadly be categorised in experimental or observational studies'*, and that *'the most reliable evidence about the relative treatment effects of a technology is obtained from experimental studies with high internal and external validity. For an assessment of internal validity, the different types of study design can be ranked according to design features that affect their validity for estimating relative treatment effect, ranging from RCTs to uncontrolled observational studies*.' However, for the purposes of MAFEIP, it is important to note that the innovative character of many technologies under assessment, and the vast amount of interventions to assess is likely to limit both the availability of high quality evidence to populate the tool and the time available to source and analyse this evidence. Nevertheless, we ought to find a pragmatic solution to populate the tool with data even in the absence of high quality evidence. We have therefore developed an algorithm for populating the tool with relevant data that aims to balance out the general desire for high quality information with good internal and external validity and the limiting factors that may require a more pragmatic approach. Our approach is illustrated in Figure 5 below.

The proposed algorithm starts off with an assessment of the data provided by commitments. It is important to note that research methods may differ across commitments, meaning that if an estimate is available that could be used to populate the tool, a brief assessment of the internal and external validity of the data provided may be indicated. On the broadest level, this could mean that we make a crude judgement based on the nature of the data reported before accepting estimates from commitments (e.g. have experimental methods been used or is the data based on a less robust observational design with or even without a control group). As a second step, we may then briefly screen the related RCT literature to check whether the same intervention has previously been evaluated in an experimental study so that estimates of treatment effect could be confirmed or

elicited and applied to the tool. If this is not the case, we may use published results from observational studies, if available. However, as mentioned above, the innovative character of many interventions within the EIP on AHA may result in a generally low evidence base, so that the subsequent step would be to use estimates from similar commitments, if available. This may require combining information available from different commitments which focus on similar interventions, patient subgroups, comparators and outcomes assessed, and this approach is predicated on the assumption that parameter inputs may be transferrable between those commitments. Further, if neither the published literature, nor similar commitments within the EIP on AHA provide useful estimates to populate the tool, the next step would be to ask experts for their judgement. If expert judgement does not help to populate the tool either, we may consider alternative secondary indicators to link up surrogate health outcomes to HRQoL and disease specific mortality (which will also briefly be discussed in Section 3.3 below). The idea behind using alternative indicators is that there may be different ways to extrapolate health outcomes from one intervention towards the EIP on AHA objectives, and if data availability for a particular intervention does not permit a specific route, alternative routes could be checked before excluding the intervention from the quantitative assessment within MAFEIP. However, this option also needs to be balanced against the additional analytic resources that may be required to conceptualise respective model extensions and populate them with relevant data (see subsequent section for more details). This is why the respective boxes in Figure 5 are shaded in grey and dashed lines have been used to underline the rather hypothetical nature of this option. If none of the above works, the commitment has to be excluded from the described quantitative modelling exercise.

It needs to be highlighted that the above algorithm constitutes a 'best case scenario' which reflects our aim to provide a monitoring framework that is, a priori, as inclusive as possible as we cannot yet foresee the amount or quality of data that is likely to become available in the future. Further adjustments may therefore be indicated once we know more about data availability within MAFEIP, and this could require scaling down the whole exercise to a more manageable level. For instance, literature searches for published evidence could be reduced to a minimum, or expert judgement could be used earlier in the process. Potential model extensions may further increase the required workload so that this option should be considered with caution.

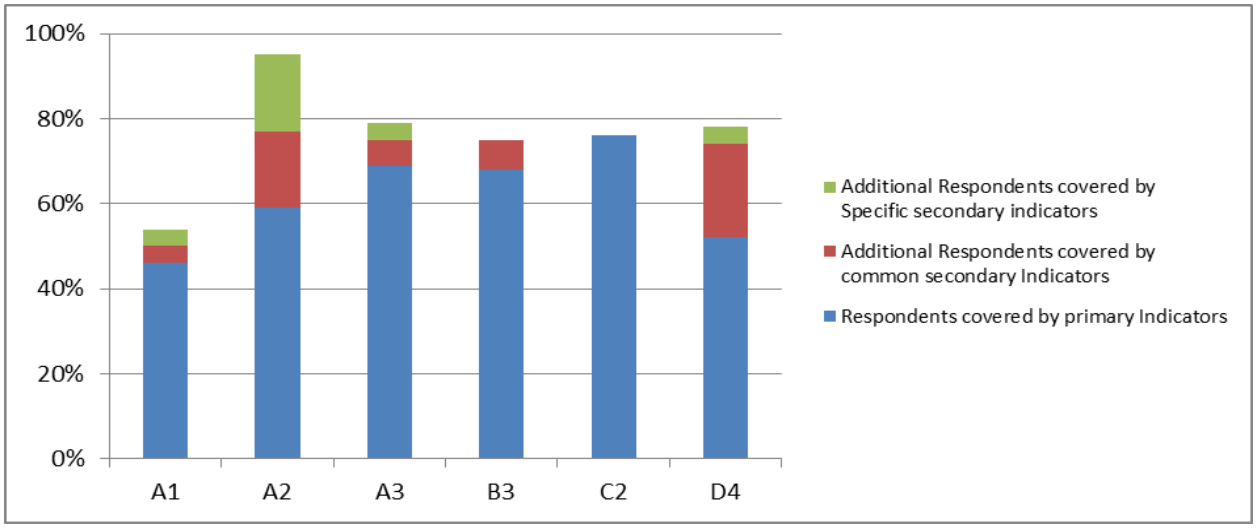
Figure 5: Algorithm to populate tool with relevant data on treatment effects



3.3 Use of secondary outcome indicators within MAFEIP

The analysis of the survey data on outcome indicators reported in the MAFEIP Second report on outcome indicators led to a number of important insights for conceptualising a quantitative modelling framework for MAFEIP. First of all, it showed that primary indicators for the QoL column (HRQoL and mortality) are likely to become available for a number of commitments participating in the EIP on AHA. Whilst primary indicators covered 63% of respondents answering this survey, common secondary indicators add 9% to the overall potential coverage, and specific secondary indicators add another 4% to overall coverage of respondents. This is also displayed in Figure 6 below.

Figure 6: Commitments potentially covered by each type of indicator per Action Group



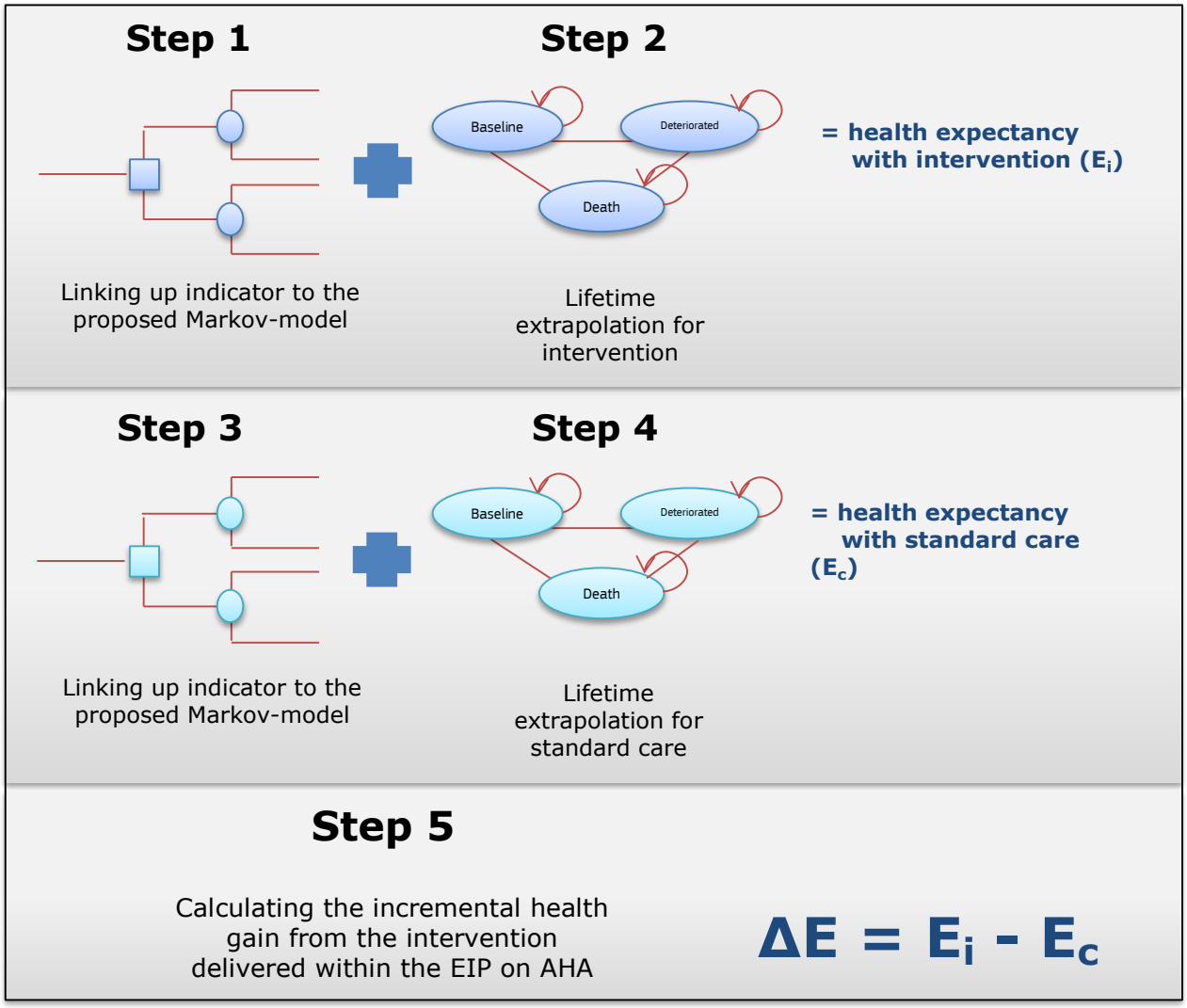
Source: Stakeholder survey, MAFEIP Second report on outcome indicators

As already mentioned in Section 3.1 , the generic tool which we aim to develop for the purposes of MAFEIP allows drawing from data related to both primary and secondary outcome indicators. More precisely, our basic model rests on three health states, two 'alive'-states ('baseline health' and 'deteriorated health') and a 'death' state respectively. The model is built around the two primary outcome indicators (HRQoL and mortality), but adapting the alive states to different target cohorts provides the opportunity to draw upon data from secondary outcome indicators which can be utilised to define mutually exclusive health states for the model. For instance, 'baseline health' could describe a particular patient cohort before the occurrence of a fall, whilst 'deteriorated health' describes the health and resource outcomes for the same population after a fall has happened. The transition between both states represents the likelihood of a fall in the target population. Likewise, different instruments which aim to assess the progression of a particular disease (such as, for instance, the Groningen Frailty Index [Steverink et al., 2001], or the Clinical COPD questionnaire [van der Molen et al., 2003]) could be used to define the baseline and deteriorated states of the model respectively for that disease, and HRQoL data from a generic instrument such as the EQ-5D could then be used to weight the respective health states of the model.

While this means that the proposed tool for MAFEIP can accommodate data related to various (primary and secondary) outcome indicators, there are a few exceptions which may require extensions to the basic model. These exceptions relate to outcome data from certain secondary outcome indicators which cannot be used directly to define respective health states for the model. In other words, an additional step would be required to link those outcome indicators to the general model in order to extrapolate respective outcomes to the EIP on AHA health objectives. We will use the example of the 'physical activity' indicator to describe a potential model extension for that particular secondary outcome indicator and will also briefly mention further possible extensions for 'risk factors' and

'adherence to treatment'. If data became available in the future on any of those indicators justifying their use, more work would be required to fully conceptualise respective model extensions to our overall framework and the feasibility of doing so would have to be assessed in light of available resources.

Figure 7: Extending the basic model for certain secondary outcome indicators

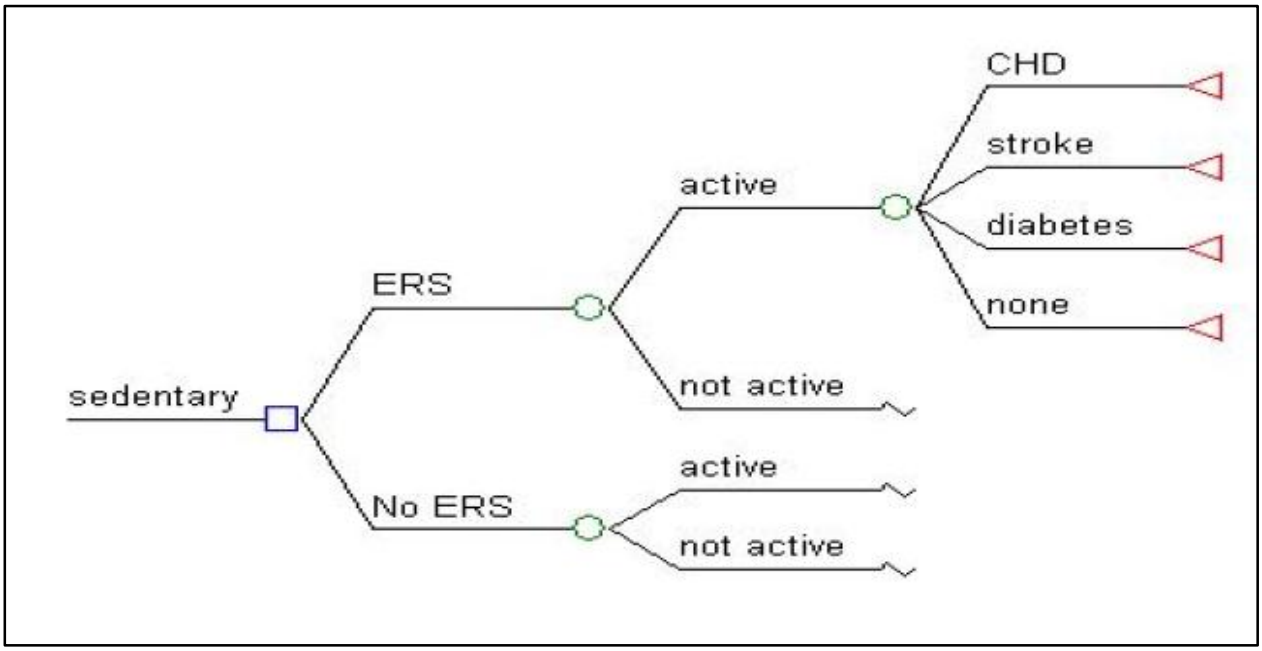


To include such secondary indicators into the modelling framework for MAFEIP, it is necessary to go beyond the basic three-stage Markov-cohort model. One way to do so is to combine a decision tree with the proposed Markov cohort model [e.g. Briggs et al., 2006], as displayed in Figure 7 above. Although decision tree models and Markov models differ in a number of aspects, they are by no means incompatible. As Briggs et al. (2006) specify, Markov models may rather be seen as 'just a form of recursive decision tree'. Hence, we may also be able to use both forms of modelling jointly.

In the context of MAFEIP, we could, for instance, first use a decision tree model to characterise the behavioural outcomes of a physical activity intervention delivered within an EIP on AHA commitment. The purpose of the decision tree model would be to estimate, both for the intervention and the standard care scenario, the proportion of the patient cohort that changes their physical activity behaviour, and subsequent changes in the probability to experience an adverse health event (such as CHD, stroke, or type 2 diabetes). This would ultimately impact on the proportions of the patient cohort in each state of the Markov model. The purpose of the Markov-model would then be to estimate the quality adjusted life expectancy with and without the intervention. The existing health economic

evaluation literature could be used to help further conceptualise respective model extensions. For instance, NICE (2006) assessed the cost-effectiveness of physical activity interventions using a decision tree in order to model the impact of a change in physical activity on the onset of coronary heart disease (CHD), stroke, type 2 diabetes and colon cancer; and the model was used to estimate the cost and quality adjusted life expectancy from the onset of these conditions. Subsequently, Anokye et al (2011) built a very similar model, though they did not link behavioural change to the likelihood of developing colon cancer as they found more robust evidence for CHD, type II diabetes and stroke (Figure 8).

Figure 8: Decision tree model for physical activity by Anokye et al (2011)



In the study by Anokye et al (2011), the authors considered a cohort of sedentary individuals (aged between 40 and 60) who are exposed to exercise referral schemes compared to a control group not exposed to the intervention. Effectiveness of the intervention was measured as the probability of moving from a sedentary to an active state based on an intention to treat analysis and data was obtained from a meta-analysis as part of a systematic literature review on the effectiveness of ERS schemes. Being active was defined as doing at least 90 to 150 minutes of moderate intensity physical activity per week. Sedentary patients were hence defined as those with activity levels below this threshold. A further key assumption (for which the authors could not provide sufficient scientific evidence though), was that the change in activity levels would last long enough to translate into future health benefits in terms of a reduced risk of developing a disease. For the purposes of MAFEIP we could, for instance, adapt the model in Figure 8 and populate it with behavioural change data obtained from commitments participating in the EIP on AHA. The purpose of the decision tree-model would be to estimate the likelihood of an adverse health event (CHD, stroke, diabetes) for individuals who could either remain inactive or become active following the intervention, and this would feed into the Markov model that rests on different transition probabilities between health states for active and inactive patient cohorts.

Respective model extensions could also be developed for other secondary indicators. For instance, risk factors (such as blood cholesterol, glucose levels, diabetes status or obesity) could be used to estimate the future risk of a CHD-event, and the subsequent probabilities to move from the 'baseline' state into the 'deteriorated health' or 'death' state of the Markov model. This could be done using validated risk stratification tools such as the Framingham risk equation [D'Agostino et al., 2008] or the Q-RISK

instrument [Hippisley-Cox et al., 2007]. For adherence to interventions, we could assume full treatment effectiveness only for the proportion of people who adhere to the intervention, and any increase in this proportion would subsequently result in better health outcomes [e.g. Greving et al., 2011]. However, as mentioned above, our intention here is to only present a conceptual framework that allows for model extensions to incorporate additional secondary indicators into MAFEIP, if necessary, and not (yet) to fully develop the respective model extensions. Future work within MAFEIP may be dedicated to develop such models if data became available that would justify their use; however, this decision should be made in the light of the additional coverage of the framework that a particular model extension may help achieve and the analytic resources available to develop them.

3.4 Adapting the model to estimate impact on the sustainability of health systems

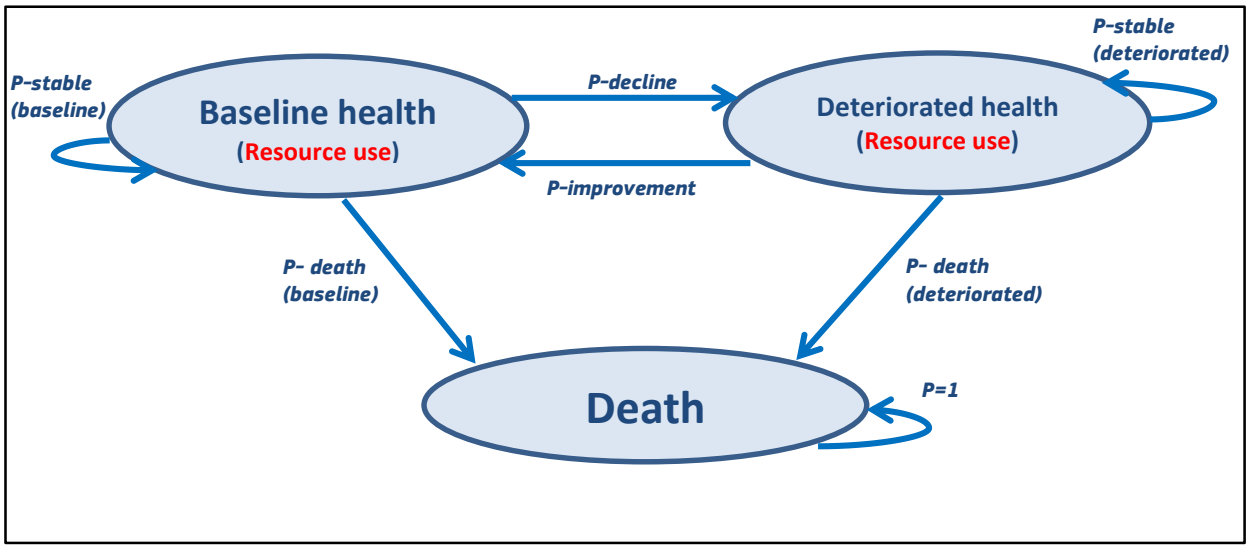
A key concern for most countries across the European Union is an ever rising demand for health and care services, resulting in constantly increasing pressure on health and care systems. Therefore, assessing how interventions delivered by the commitments participating in the EIP on AHA may impact on the sustainability of health and care systems through their respective impact on health and care expenditure across the EU is a key objective of MAFEIP. This includes both the cost associated with implementing (and scaling up) the health and social care models proposed within different commitments and Action Groups, and the potential cost savings from the respective interventions that may arise, for instance, from avoided care episodes in the future, or the shift from more to less resource intensive care alternatives. In the MAFEIP first and second reports on outcome indicators we have proposed collecting information on the quantities of resources related to a particular intervention compared to a standard care alternative to assess the EIP on AHA impact on the sustainability of health and care systems. We further proposed to value such resources in monetary units using national average unit cost data, which would then enable quantification of an intervention's impact on respective health and care expenditures.

We believe that our modelling proposal for the quality of life column bears a particular advantage as it may also be used to assess the EIP on AHA impact on the sustainability of health and care systems through its estimated impact on health and care expenditures. More precisely, we could utilise Markov based cohort models which were previously developed and populated with data for the assessment of health impact in order to estimate interventions health and care resource use, and their consequent impact on health and care expenditure in a particular jurisdiction. Using the same models for both columns of the monitoring framework ensures a more consistent approach to the assessment of impact and a much more efficient use of analytic resources within the MAFEIP project.
Hence, our approach would rest on two key principles:

- First, we may use the Markov models originally developed and populated with data for monitoring health impact within MAFEIP in order to estimate the *'prevalence'* of the different *'alive-states'* relevant for a particular intervention within a certain population [e.g. Mar et al., 2008].
- Secondly, by replacing HRQoL weights with resource use data from commitments that could be weighted with national average unit cost, we can then estimate the incremental impact of a particular intervention over standard care on health and care expenditure.

All other aspects of the modelling strategy as outlined above within the context of health impact assessment would hold for the estimation of impact on the sustainability of health and care systems through the estimated impact on health and care expenditures, in particular the incremental approach that compares the commitments intervention with a respective care alternative. A schematic presentation of the adapted model is displayed in Figure 9 below.

Figure 9: Adapting the model to assess impact on the sustainability of health and care systems



Though the purpose of this document is to develop a conceptual framework rather than to fully specify a model for assessing the impact of interventions delivered within the EIP on AHA on the sustainability of health and care systems through their respective impact on health and care expenditure, it is worthwhile mentioning that there is a considerable body of literature that could be used to further inform the development of the methods for carrying out such an assessment within MAFEIP. Apart from the research that is generally involved with the economic evaluation of healthcare technologies (in particular their cost-effectiveness), both national regulatory agencies (such as the National Institute for Health and Care Excellence (NICE) in England and Wales) and the related scientific literature offer guidelines on the assessment of interventions impact on health and care resources in the context of their use within particular healthcare markets. Good practice guidelines also exist for *'budget impact analysis'*, such as those published by the task force of the International Society for Pharmaco-economics and Outcomes Research [Mauskopf et al., 2007]. The potential use of Markov Models has been explicitly mentioned in this context.

As with the estimation of health impact, however, the key challenge within MAFEIP is its wide scope, with hundreds of commitments in six thematic Action Groups implementing a vast variety of interventions across all 28 member countries (and sometimes even beyond). This calls for a model that balances out the trade-off between context specificity and existing constraints inherent to the analytic resources available. With the generic Markov model implemented as a web-based tool, we believe that such an assessment may be possible, especially as *'the purpose is not to produce exact estimates of the [budget] consequences of an intervention, but to provide a valid computing framework (a model) that allows users to understand the relation between the characteristics of their setting and the possible budget consequences of a new health technology'* [Mauskopf et al., 2007]. This evaluation should include not just the cost arising from implementing and running the new care alternative, but also a) the potential impact on related future health and care resource use that may differ between standard care and the new intervention scenario (for instance through avoided future care episodes) and b) the shift from one health and care alternative to another (e.g. from secondary to primary care). In general, the resource use considered should be *'relevant to the health condition and intervention of interest over the chosen time horizon'* [Mauskopf et al., 2007].

However, it is also important to ensure that the analysis of resource utilisation does not exceed reasonable efforts. As a general rule, it may not be worthwhile to invest huge efforts to collect resource use information whose impact on respective expenditures is relatively small. Hence, resources which are consumed in small quantities and are characterised by low unit cost may not necessarily be reported for the purposes of monitoring the impact of the EIP on AHA on the sustainability of health

and care systems [e.g. Drummond et al., 2005]. Also, resource use shall be reported only for those resource items that are likely to differ in their consumption between standard care and the intervention delivered within the EIP on AHA within each state of the model. The idea is that, by choosing an incremental approach, those resources whose consumption does not change through the intervention would be the same for standard care as for the intervention scenario [e.g. Drummond et al. 2005].

Another question to answer is that of the economic perspective, or who bears the costs of the resources consumed / saved as a result of the intervention. A strict healthcare perspective (i.e. not taking into account cost in related sectors such as the social care system) may be a too restrictive viewpoint for the purposes of MAFEIP. We propose, at a minimum, a "health and care" perspective that simultaneously looks into the impact of the EIP on AHA on both health and care systems. As the boundaries between health and care systems may also differ between different countries, choosing a perspective that considers both may improve the comparability of results across different geographic settings. For instance, whilst some care related services are commonly being delivered by community nurses in a UK setting (and therefore not considered from a strict NHS perspective), the same services may fall within the healthcare system in other countries where community nurses do not exist. However, an even broader perspective may also be indicated for several reasons. For some commitments or Action Groups, costs may be more restricted to health and care systems than for others. Only considering health and care resources may therefore bias results towards or away from particular Action Groups if substantial investments or savings lie outside the health and care systems. This may be particularly the case for Action Group D4. Secondly, assuming that the overall impact of EIP on AHA activities is positive, it needs to be considered whether this result would come at the cost of increased pressure on the budgets available in other sectors, which may ultimately require a societal perspective. Note that we do not see this to be in conflict with the good practices for budget impact analysis as suggested by Mauskopf et al. (2007), as it follows the recommendation to take the perspective of the respective *'budget holder'*. As multiple 'budget holders' may be affected by EIP on AHA interventions across countries, a societal perspective may be the only valid option.

4. Issues for discussion

The purpose of this report was to propose a quantitative modelling framework for linking outcomes of the EIP on AHA policy initiative to its objectives related to quality of life and sustainability of health and care systems. Whilst we believe to have provided a robust proposal that is technically feasible, it needs to be highlighted, however, that the proposed framework will require further adjustment and fine tuning in the future to adapt it to the actual data availability. At such an early stage of the project, where both the number of commitments that will provide outcome data and the quality of the information to be received is unclear, it is not possible to tailor our proposal further to anticipate future needs. Rather, we aimed to provide a flexible framework that does not, a priori, exclude any feasible option to capture interventions' outcomes. This is reflected, for instance, in the generic model framework which builds upon data from both primary and secondary outcome indicators, and also allows including additional health states or even model extensions to accommodate for additional indicators should their use increase the flexibility and consequently the coverage of commitments through the framework. Our proposed framework also shows flexibility insofar it allows populating the model with data from secondary sources in order to assess the health and sustainability outcomes of a particular intervention in cases where this can help fill gaps in the existing data. In that context, we provided an example of algorithm that supports the choice of appropriate data sources for populating the model.

While our proposal offers the flexibility to adjust and fine tune the modelling framework to the reality of future data availability, the examples above demonstrate, in a sense, the current 'best case' character of the proposed framework. For instance, if a large number of commitments provide suitable data, this may allow assessing the impact of interventions' outcomes for many commitments by adapting the basic model. However, at the same time analytic resources may not be sufficient to extend the framework further by modelling impact through possible model extensions as discussed in Section 3.3. Further work will therefore be required to assess the technical feasibility of developing such model extensions for linking up certain secondary indicators to the EIP on AHA objectives. With respect to populating models with data, it is also likely that this will constitute a key constraint to the future coverage of our proposal, as the diversity of the interventions delivered within the EIP on AHA and also their innovative character may result in a generally scarce and scattered evidence base. Even a crude search for input values for parameters may prove time consuming.

We therefore believe that a number of options ought to be considered in order to further adapt our modelling framework to potential future needs and requirements. The obvious default strategy for MAFEIP would be that, following the reporting of outcomes data by commitments, IPTS in collaboration with DG SANCO and DG CONNECT undertakes all subsequent steps in order to adapt a model to a particular intervention. This would involve verification of estimates provided by commitments, populating respective models with outcomes data, and closing gaps in the data through reviews of secondary data sources (such as the published literature, estimates provided by similar commitments, or expert opinion). The advantage of this approach would be that IPTS remains in full control over the validity of the data inputs, and that similar quality standards for data inputs can be fulfilled across commitments. However, the respective efforts for data sourcing, validation and input into respective models would be tremendous, and it is unlikely that IPTS could take over this task for more than a small subset of commitments participating in the EIP on AHA (it is impossible to estimate the number of interventions for which this approach could be feasible at this point as it depends on many factors, but overall coverage would most likely remain low).

With respect to implementing models and populating them with intervention specific data, we therefore propose an alternative implementation strategy for MAFEIP, which is the development of a tool that may enable stakeholders within commitments to assess health and health and care systems' sustainability impact of their respective interventions independently, with the support and guidance of IPTS. Initially, this tool should be based on the generic three-state Markov model described here, but in

order to increase its flexibility to be adapted to different contexts (which would further increase the coverage of commitments through our proposal), the optional inclusion of more than three health states should be considered for the future. The key advantage of this strategy would be to reduce the burden of data collection within MAFEIP towards a more manageable level by involving experts within commitments in the process of gathering, validating, and selecting the evidence required to populate the tool. Further, stakeholders could benefit from obtaining access to a tool that helps them evaluate the impact of their respective interventions on patient's health and the sustainability of health and care systems; and they could also improve their knowledge base with respect to decision modelling methods for health economic evaluation in general, and evaluating the impact of their respective health and care interventions in particular. IPTS could facilitate the use of the tool through dedicated workshops and through direct communication with commitments participating in the EIP on AHA. The analytic resources that could be freed through this approach could be used to further increase the flexibility of the proposed monitoring framework, for instance, through the optional inclusion of additional health states in the model, or by developing model extensions for incorporating selected secondary indicators. The main drawback of this approach, however, would be a loss of control over the quality of the data used to populate the tool, so that this strategy would require developing effective quality checks of the data. To address this potential loss of control over the quality of data used to populate the tool, sufficient support should be provided to commitments participating in the EIP on AHA. Therefore, an important activity to be envisaged for the continuation of MAFEIP would be for IPTS to facilitate the use of the tool, raise awareness about its purpose and the advantages it provides for EIP on AHA stakeholders, and to also support commitments with populating the tool with data for their own purposes.

Another area that we identified for further research within MAFEIP is the use of mapping algorithms to address the problem that commitments may use a variety of instruments to elicit information on patients HRQoL. As a consequence, HRQoL data may not be comparable across interventions and without addressing this issue, it could be impossible to aggregate health gain across the groups of interventions for which different instruments have been applied. However, as mentioned in Section 3.1, the scientific literature offers methods for 'mapping', or 'cross-walking' between different elicitation methods [e.g. Dakin, 2013], and in order to increase the generalizability of our proposed framework, we propose either calculating or adapting existing algorithms that are published in the scientific literature to 'convert' HRQoL-values measured with one instrument into equivalent values of another instrument. There are both theoretical and practical reasons why a mapping from other instruments towards the EQ-5D may be indicated. We therefore suggest to conduct more research into the use of mapping for the purposes of MAFEIP, and to further consult with external experts on this particular matter.

5. Conclusions

The purpose of this report was to propose a quantitative modelling framework for linking outcomes of the EIP on AHA policy initiative to its objectives related to quality of life and sustainability of health and care systems.

After confirming the selection of potential outcome indicators, the main challenge that we faced was the linking between these indicators on the one hand and the EIP on AHA targets of adding 2HLY to the life of European citizens and achieving a Triple Win for Europe on the other. After a thorough search of the scientific literature on potential ways of bridging the outcomes of the EIP on AHA and the above targets, we decided to take the HLY model as a starting point and adapt it to the needs of MAFEIP. The rationale for this adaptation was to make sure the resulting model could adequately capture (incremental) health impacts achieved by EIP on AHA commitments, and also to use data on indicators that are most frequently reported across the EIP on AHA by participants. As a result, we have been able to propose a generic and flexible modelling approach based on a Markov process, which allows us to link (primary and secondary) outcome indicators across interventions, commitments and geographic domains to the EIP on AHA health objective. We further demonstrated how some model extensions could be implemented to incorporate additional secondary indicators in the overall framework, although this would depend on technical feasibility and the availability of analytical resources. Another advantage of this model is that it can be adapted to estimate the impact of activities delivered within the EIP on AHA on the sustainability of health and care systems in terms of changes in health and care expenditure. Finally, we propose that this framework should be implemented as a tool that could enable stakeholders within commitments to assess the health and sustainability impact of their respective interventions independently, with the support and guidance of IPTS.

This report has therefore brought us to the point where we can actually start building a model to estimate the impact of the innovations delivered within the EIP on AHA and their contribution to the Quality of Life of European citizens and the sustainability of health and care systems. Nevertheless, this implies that relevant data from the Action Groups and commitments will become available which may not be the case immediately. This is why the issue of data availability, data collection and data sources for populating the model will play a crucial role in determining how we can best implement our approach in practice. While we believe that developing a monitoring tool would facilitate the data collection and input process, some issues will have to be resolved. We therefore deem it important to facilitate further discussions on data issues and actively promote the framework among stakeholders, commitments or Action Groups. Indeed, explaining our model and its advantages will help to ensure the acceptance of the approach proposed by all those involved. We believe this would be the best way to implement the framework successfully, and in turn estimate the real impact of the EIP on AHA.

6. References

Anokye N. K., Trueman P., Green C., et al. (2011). The cost-effectiveness of exercise referral schemes. BMC Public Health;11:954.

Anokye N. K., Lord, J., Fox-Rushby, J. (2013). Is brief advice in primary care a cost-effective way to promote physical activity? BMJ Open Sport & Exercise Medicine, online first, Dec. 18th, doi:10.1136/bjsports2013-092897

Birch S, Gafni A. (2003). Economics and the evaluation of health care programmes: generalisability of methods and implications for generalisability of results. Health Policy. 2003 May;64(2):207-19.

Boehler, C. E. H.. (2013). Mind the Gap: Geographic Transferability of Economic Evaluations in Health – a thesis submitted for the degree of Doctor of Philosophy. Brunel University, Health Economics Research Group (HERG), London, UK. Available at: http://bura.brunel.ac.uk/handle/2438/7170

Boehler C. E. H., Lord J. (2015) Mind the Gap! A Multilevel Analysis of Factors Related to Variation in Published Cost-Effectiveness Estimates within and between Countries. Medical Decision Making (ahead of print) doi: 10.1177/0272989X15579173

Brazier, John & Roberts, Jennifer & Deverill, Mark, (2002). "The estimation of a preference-based measure of health from the SF-36," Journal of Health Economics, Elsevier, vol. 21(2), pages 271-292, March.

Brazier J, Yang Y, Tsuchiya A. Rowan DL (2010). A review of studies mapping (or cross walking) from non-preference based measures of health to generic preference-based measures, European Journal of Health Economics 11:215-25

Briggs, A., Sculpher, M. (1998). An introduction to Markov modelling for economic evaluation (1998) PharmacoEconomics, 13 (4), pp. 397-409.

Briggs, A, Claxton, K, Sculpher, M (2006). Decision modelling for health economic evaluation. Oxford University Press, Oxford, UK

Brooks R, Rabin R, de Charro F, (eds) (2003). The Measurement and Valuation of Health Status Using EQ-5D: A European Perspective: Evidence from the EuroQol BIO MED Research Programme. Rotterdam: Kluwer Academic Publishers.

Craven, MP, Allsop, MJ, Morgan, SP, Martin, JL (2012). Engaging with economic evaluation methods: insights from small and medium enterprises in the UK medical devices industry after training workshops. Health Research Policy and Systems, 10 (September), 29. Available at: http://www.health-policy-systems.com/content/10/1/29

Crove, JA, Craven, MP, Lu, B, Morgan, SP (2010) Experiences of using a spreadsheet decision-tree tool with non-professional health economists. Proceedings of the 7th annual conference of the International Society for Health Technology Assessment (HTAI 2010)

D'Agostino, R.B.; Sr., Vasan, R.S., Pencina, M.J., Wolf, P.A., Cobain, M., Massaro, J.M., Kannel, W.B. (2008). "General cardiovascular risk profile for use in primary care: the Framingham Heart Study". Circulation 117 (6): 743–753.

Dakin, H (2013). Review of studies mapping from quality of life or clinical measures to EQ-5D: an online database. Health and Quality of Life Outcomes, vol. 11, 151.

Drummond, MF (1994). Comparing cost-effectiveness across countries. The model of acid related disease. Pharmacoeconomics, vol.5, pp. 60-67

Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL (2005). Methods for economic health evaluation of health care programmes. 3rd ed. Oxford: Oxford University Press

Eldessouki R, Smith MD (2012). Health Care System Information Sharing: A Step Toward Better Health Globally. Value Health Regional Issues; 1:118-129

European Commission (2011): Strategic Implementation Plan for the European Innovation Partnership on Active and Healthy Ageing Steering Group Working Document, Final text adopted by the Steering Group on 7/11/11, Strategic Plan. Available at: http://ec.europa.eu/research/innovation-union/pdf/active-healthy-ageing/steering-group/implementation_plan.pdf#view=fit&pagemode=none

EUROSTAT Glossary (2014): Healthy life years statistics. Available at http://epp.eurostat.ec.europa.eu/statistics_explained/index.php/Healthy_life_years_statistics (last accessed: 04th July 2014)

EUROSTAT, Population Data. (2014) Available at: http://epp.eurostat.ec.europa.eu/portal/page/portal/population/introduction (last accessed: 04th July 2014)

EUROSTAT: EU SILC (2014) ' EU statistics on income and living conditions'. Available at: http://epp.eurostat.ec.europa.eu/portal/page/portal/microdata/eu_silc (last accessed: 4th July, 2014)

Goeree, R., Burke, N., O'Reilly, D., Manca, A., Blackhouse, G. & Tarride, J. (2007). Transferability of economic evaluations: approaches and factors to consider when using results from one geographic area for another. Current Medical Research and Opinion, 23(4), pp. 671-682.

Greving, J., Visseren, F., De Wit, G. & Algra, A. (2011). Statin treatment for primary prevention of vascular disease: whom to treat? Cost-effectiveness analysis. BMJ (Clinical research ed.), 342, pp. d1672.

Haiycox, A, Dubois, D, Butterworth, M (1998). Customising an international disease management model to the needs of individual countries: Application to upper gastrointestinal disease. PharmacoEconomics, 14 (SUPPL. 2), pp. 39-56.

Hippisley-Cox J et al. (2007). Derivation and validation of QRISK, a new cardiovascular disease risk score for the United Kingdom: prospective open cohort study. BMJ 335: 136

Hunt SM, McKenna SP, McEwen J, Williams J, Papp E. (1981). The Nottingham Health Profile: subjective health status and medical consultations. Soc Sci Med A.;15(3 Pt 1):221–229.

International Society for Pharmacoeconomics and Outcomes Research (ISPOR). ISPOR Pharmacoeconomic Guidelines Around The World. Available at: http://www.ispor.org/PEguidelines/index.asp. [July, 067 2014].

Jagger, C, Cox, B, Le Roy, S & the EHEMU team (2006). Health Expectancy Calculation by the Sullivan Method. Third Edition. EHEMU Technical Report September 2006

Jagger, C, Gillies, C, Moscone, F, Cambois, E, Van Oyen, H, Nusselder, W, Robine, JM & the EHLEIS team (2008). Inequalities in healthy life years in the 25 countries of the European Union in 2005: a cross-national meta-regression analysis. The Lancet, 372: 2124-2131

Longworth L, Rowen DL (2013). Mapping to obtain EQ-5D utility values for use in NICE health technology assessments. Value In Health 16(1)202 -10

Longworth L, Yang Y, Young T, Brazier J, Rowen D, Tsuchiya A, et al (2014). Use of generic and condition-specific measures in NICE decision-making. Health Technology Assessment , Volume 18, number 9.
Mar, J, Sainz-Ezkerra, M, Miranda-Serrano, E (2008). Calculation of prevalence with Markov models: budget impact analysis of thrombolysis for stroke. Medical Decision Making, vol. 28, pp 481-490

Mauskopf, JA, Sullivan, SD, Annemans, L, Caro, J, Mullins, CD, Nuijten, M, Orlewska, E, Watkins, J, Trueman, P (2007). Principles of good practice for budget impact analysis: report of the ISPOR Task Force on good research practices – budget impact analysis. Value in Health, vol. 10, pp. 336 – 347

NICE (2006) Modelling the cost-effectiveness of physical activity interventions. National Institute for Health and Care Excellence. Available at: https://www.nice.org.uk/guidance/ph2/documents/physical-activity-economics-modelling-report2

NICE (2013) Guide to the Methods of Technology Appraisal. National Institute for Health and Care Excellence. London, 2013. Available at: http://publications.nice.org.uk/guide-to-the-methods-of-technology-appraisal-2013-pmg9/foreword

O'Brien (1996). Economic evaluation of Pharmaceuticals – Frankenstein's Monster or Vampire of Trials. Medical Care, vol. 34, No. 12, suppl. Pp.DS99-DS108

Robine J.M, Romieu I. and Cambois E. (1999) Health expectancy indicators. Bulletin of the World Health Organization, 1999, 77 (2).

Sculpher, M.J., Pang, F.S., Manca, A., Drummond, M.F., Golder, S., Urdahl, H., Davies, L.M. & Eastwood, A. (2004). Generalisability in economic evaluation studies in healthcare: a review and case studies. Health Technology Assessment, 8(49), pp. iii-iv, 1-192.

Sintonen H. (2001). The 15D instrument of health related quality of life: properties and applications. Ann Med 2001, 33:328-336.

Steverink N., Slaets J.P.J., Schuurmans H., van Lis M. (2001) Measuring frailty: Development and testing of the Groningen Frailty Indicator (GFI).Gerontologist, 41(special issue 1):236-237.

Stinnett, A.A. & Mullahy, J. (1998). Net health benefits: A new framework for the analysis of uncertainty in cost-effectiveness analysis. Medical Decision Making, 18(2 SUPPL.).

van der Molen T., Willemse B.W., Schokker S., ten Hacken N.H., Postma D.S., Juniper E.F. (2003) Development, validity and responsiveness of the Clinical COPD Questionnaire. Health Qual Life Out; 1: 13.

Ware JE, Kosinski M, Keller SD (1996). A 12-item short form health survey (SF-12): construction of scales and preliminary tests of reliability and validity. Med Care; 32: 220–33.

Ware JE, Snow KK, Kosinski M et al.(2000). SF-36 Health Survey: Manual and Interpretation Guide. Lincoln, RI: QualityMetric Incorporated 1993.

Weinstein, MC, Torrance, G, McGuire A (2009). QALYs: The basics. Value in health, vol.9, suppl. 1, pp. S5-S9

Weinstein MC (1988). A QALY is a QALY is a QALY – or is it? Journal of Health Economics, vol. 7, pp-289-290

7. Annexes

Appendix 1: Systematic search strategy for identifying papers on the bridging between QALYs and HLYs

Appendix 1.1: database search strategies

SCOPUS search strategy:

Performed on March, 18th 2014 (covers EMBASE, MEDLINE and SCIENCE DIRECT)

Nr	Search terms	Hits
1	ALL("quality adjusted life year" OR "quality adjusted life years" OR QALY OR QALYs)	17,628
2	ALL("disability adjusted life year" OR "disability adjusted life years" OR DALY OR DALYs)	301,814
3	ALL("Healthy life year" OR "Healthy life years" OR HLY OR HLYs)	1,957
4	#1 AND #3	46
5	#2 AND #3	92
6	#4 AND #5	30
7	**#4 OR #5**	**108**

PubMed search strategy:

Performed on March, 19th 2014

Nr	Search terms	Hits
1	ALL("quality adjusted life year" OR "quality adjusted life years" OR QALY OR QALYs)	9986
2	ALL("disability adjusted life year" OR "disability adjusted life years" OR DALY OR DALYs)	9854
3	ALL("Healthy life year" OR "Healthy life years" OR HLY OR HLYs)	776
4	#1 AND #3	12
5	#2 AND #3	16
6	#4 AND #5	4
7	**#4 OR #5**	**25**

Cochrane Library search strategy:

Performed on March, 24th 2014

(Quality adjusted life year OR quality adjusted life years OR QALY OR QALYs or Disability adjusted life year OR Disability adjusted life years OR DALY OR DALYs) AND (Healthy life year OR Healthy life years OR HLY OR HLYs)

Database	Hits
Cochrane Database of Systematic Reviews (CDSR)	1
Cochrane Central Register of Controlled Trials	25
Cochrane Methodology Register	3
Database of Abstracts of Reviews of Effects (DARE)	0
Health Technology Assessment Database (HTA)	0
NHS Economic Evaluation Database (NHS EED)	8

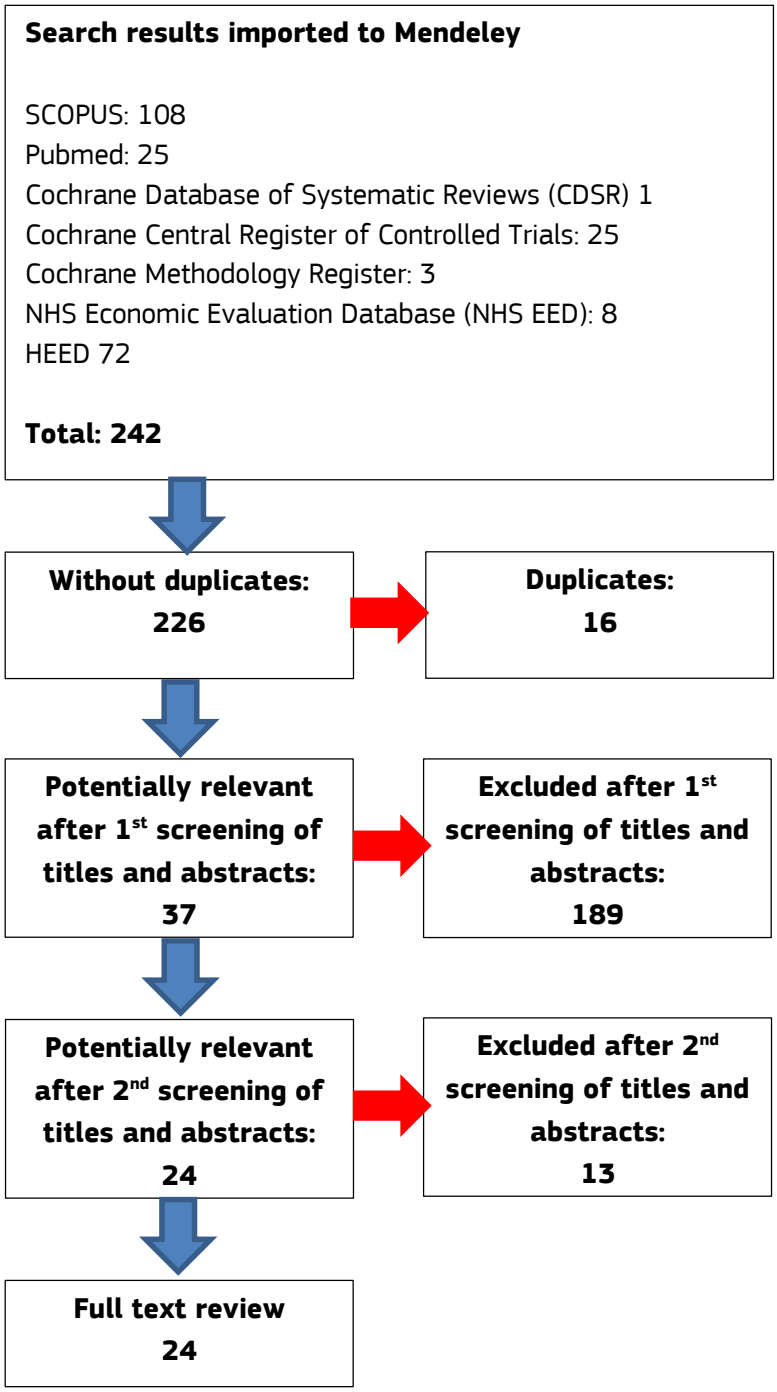
HEED search strategy (Health Economics Evaluation Database)

Performed on March, 24th 2014

(Quality adjusted life year OR quality adjusted life years OR QALY OR QALYs or Disability adjusted life year OR Disability adjusted life years OR DALY OR DALYs) AND (Healthy life year OR Healthy life years OR HLY OR HLYs) in Abstract

No of hits: 72

Appendix 1.2: Managing references



Appendix 1.3: Papers under full-text review

Canning, D. (2013). Axiomatic foundations for cost-effectiveness analysis. *Health economics*, *22*(12), 1405–16. doi:10.1002/hec.2889

Cuijpers, P., Smit, F., Unger, F., Stikkelbroek, Y., Ten Have, M., & De Graaf, R. (2011). The disease burden of childhood adversities in adults: a population-based study. *Child abuse & neglect*, *35*(11), 937–45. doi:10.1016/j.chiabu.2011.06.005

De Hollander, A. E. M., Melse, J. M., Lebret, E., & Kramers, P. G. N. (1999). An aggregate public health indicator to represent the impact of multiple environmental exposures. *Epidemiology* (Vol. 10, pp. 606–617). Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-0032783878&partnerID=tZOtx3y1

Eralp, M. N., Scholtes, S., Martell, G., Winter, R., & Exley, A. R. (2012). Screening of healthcare workers for tuberculosis: development and validation of a new health economic model to inform practice. *BMJ open*, *2*(2), e000630. doi:10.1136/bmjopen-2011-000630

Gennaro, V., Ghirga, G., & Corradi, L. (2012). In Italy, healthy life expectancy drop dramatically: from 2004 to 2008 there was a 10 years drop among newborn girls. *Italian journal of pediatrics*, *38*(1), 19. doi:10.1186/1824-7288-38-19

Gromulska, L., Wysocki, M. J., & Goryński, P. (2008). Lata przezyte w zdrowiu (healthy life years, HLY)--zalecany przez Unie Europejska syntetyczny wskaźnik sytuacji zdrowotnej ludności. *Przegląd epidemiologiczny*, *62*(4), 811–820. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/19209744

Hyder, A. A., & Morrow, R. H. (2000). Applying burden of disease methods in developing countries: A case study from Pakistan. *American Journal of Public Health*, *90*(8), 1235–1240. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-0033916424&partnerID=tZOtx3y1

Hyder, Adnan A., & Morrow, R. H. (1999). Steady state assumptions in DALYs: Effect on estimates of HIV impact. *Journal of Epidemiology and Community Health*, *53*(1), 43–45. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-0032897390&partnerID=tZOtx3y1

Hyder, Adnan A., Rotllant, G., & Morrow, R. H. (1998). Measuring the burden of disease: Healthy life-years. *American Journal of Public Health*, *88*(2), 196–202. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-0031892566&partnerID=tZOtx3y1

Iburg, K. M., & Kamper-Jørgensen, F. (2002). Summary measures of population health: An overview. *Danish Medical Bulletin*, *49*(3), 256–259. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-4243288693&partnerID=tZOtx3y1

Iburg, K. M., & Kamper-Jørgensen, F. (2003). Nyere mål for en nations sundhedstilstand - En oversigt. *Ugeskrift for Laeger*. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-0042198883&partnerID=tZOtx3y1

Knol, A., Van Kamp, I., De Hollander, G., Van Kempen, E., Schram, D., Houthuijs, D., & Lebret, E. (2009). DALYs are the answer, but what was the question? *38th International Congress and Exposition on Noise Control Engineering 2009, INTER-NOISE 2009* (Vol. 4, pp. 2458–2468). Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-84870047481&partnerID=tZOtx3y1

Liang, L. Y., Huang, S. M., & Lan, C. F. (2005). Development and prospects for summary measures of population health. *Taiwan Journal of Public Health*. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-24744457939&partnerID=tZOtx3y1

Manuel, D G. (2002). Measuring the health burden of chronic disease and injury using health adjusted life expectancy and the Health Utilities Index. *Journal of Epidemiology & Community Health*, *56*(11), 843–850. doi:10.1136/jech.56.11.843

Manuel, D. G., & Schultz, S. E. (2004). Health-Related Quality of Life and Health-Adjusted Life Expectancy of People With Diabetes in Ontario, Canada, 1996-1997. *Diabetes Care*, *27*(2), 407–414. doi:10.2337/diacare.27.2.407

Mullahy, J. (2001). Live long, live well: quantifying the health of heterogeneous populations. *Health economics*, *10*(5), 429–40. doi:10.1002/hec.633

Murray, C. J. L., Salomon, J. A., & Mathers, C. (2000). A critical examination of summary measures of population health. *Bulletin of the World Health Organization*. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-0033837240&partnerID=tZ0tx3y1

Murray, C. J. L., Vos, T., Lozano, R., Naghavi, M., Flaxman, A. D., Michaud, C., Ezzati, M., et al. (2012). Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*, *380*(9859), 2197–223. doi:10.1016/S0140-6736(12)61689-4

Pang, L., Xu, Q., & Jin, S. (1999). [Healthy life years(HeaLY): a new indicator for measuring the burden of diseases]. *Wei sheng yan jiu = Journal of hygiene research*, *28*(2), 125–7. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-1842835639&partnerID=tZ0tx3y1

Pinheiro, P., & Krämer, A. (2009). Calculation of health expectancies with administrative data for North Rhine-Westphalia, a Federal State of Germany, 1999-2005. *Population health metrics*, *7*, 4. doi:10.1186/1478-7954-7-4

Rehm, J., & Frick, U. (2010). Valuation of health states in the US study to establish disability weights: lessons from the literature. *International journal of methods in psychiatric research*, *19*(1), 18–33. doi:10.1002/mpr.300

European Commission
EUR 27412 EN – Joint Research Centre – Institute for Prospective Technological Studies

Title: **Monitoring and Assessment Framework for the European Innovation Partnership on Active and Healthy Ageing (MAFEIP) Conceptual description of the Monitoring and Assessment Framework for the EIP on AHA**

Authors: Christian Boehler, Fabienne Abadie

Luxembourg: Publications Office of the European Union
2015 – 44 pp. – 21.0 x 29.7 cm

EUR – Scientific and Technical Research series – ISSN 1831-9424 (online)
ISBN 978-92-79-50574-4 (PDF)
doi:10.2791/290381

LF-NA-27412-EN-N

JRC Mission

As the Commission's in-house science service, the Joint Research Centre's mission is to provide EU policies with independent, evidence-based scientific and technical support throughout the whole policy cycle.

Working in close cooperation with policy Directorates-General, the JRC addresses key societal challenges while stimulating innovation through developing new methods, tools and standards, and sharing its know-how with the Member States, the scientific community and international partners.

Serving society
Stimulating innovation
Supporting legislation

doi:10.2791/290381

ISBN 978-92-79-50574-4

Publications Office