Ref. Ares(2017)4176608 - 25/08/2017

Deliverable

14.2 Methodologies for the measurement and valuation of costs and benefits of COMPARE

Version: 1

Due: Month 30

Completed: Month 30

Contributing Partners: EUR, CIVIC

Contributing authors EUR: Meg Perry-Duxbury, Job van Exel, Klas Kellerborg, Pieter van Baal, Werner Brouwer

Contributing authors Civic Consulting: Frank Alleweldt, Senda Kara, Camille Salinier, Kris Best

Contents

Deliverable description 4

1. Introduction 5

2. Overview of steps taken 5

3. Methodology for measuring and valuing elements of the system using case studies 6

3.1 Objectives and scope of the case studies 6

3.1.1 Definition of the system subject to the cost-effectiveness estimation 6

3.1.2 Scope of activities to be assessed 9

3.1.3 Perspective of the analysis 11

3.1.3 Time period covered and geographic scope 12

3.2 Selection of case studies 13

3.2.1 Criteria for case study selection 13

3.2.2 Description of candidate case studies 13

3.3 Assessment of costs and benefits 17

3.3.1 Evaluation of costs 17

3.3.2 Evaluation of benefits 21

3.3.3 Cost-effectiveness estimation 22

4. Measurement and valuation of elements related to the wider cost effectiveness framework 24

4.1 Value of safety 24

4.1.1 Theoretical Background 24

4.1.2 Methods 26

4.1.3 Findings 30

4.1.4 Discussion 57

5. Next steps 59

5.1 Implementation of the case studies to measure and value elements of the system.................................. 59

5.2 Implementation of the methodology to measure and value elements related to the wider framework 59

5.3 Next deliverables 59

References 60

Deliverable description

Work Package 14 aims to develop a standardised framework for estimating the cost-effectiveness of the COMPARE system and related methods and tools, including the value of safety. This second deliverable corresponds to the second and third objectives of the Work Package:

- To identify and where necessary develop state-of-the art costing methodologies for the different elements in the cost effectiveness framework;
- To develop and apply a methodology to value safety (provided through rapid identification of pathogens through COMPARE) in several countries.

This deliverable describes the steps taken so far and presents an overview of methodologies for the measurement and valuation of the elements specified in the cost-effectiveness framework, which are to be applied in a number of case studies during the next project phase.

1. Introduction

This is the second deliverable of Work Package 14, which aims to develop a standardised framework for estimating the cost-effectiveness of the COMPARE system and related methods and tools, including the value of safety. The activities of the work package are carried out jointly by WP partners Civic Consulting and Erasmus University Rotterdam (EUR).

In this deliverable, we present the work conducted 'to identify and where necessary develop state-of-the-art costing methodologies for the different elements in the [cost-effectiveness] framework' in line with Specific Objective 2 and related Task 2 of WP14, focusing on Sub-task 2.1: 'costing methodologies of elements related to system components, and selected methods and tools' and Sub-task 2.2 'Costing methodologies of elements related to wider framework'. The deliverable also outlines our approach to develop and apply a methodology to value safety in several countries (Task 4). The methodologies presented here will be refined in our subsequent work and form the basis for the related cost-effectiveness case studies in the context of Objective 4/Task 5. The document is structured as follows:

- Section 2 provides an overview of the steps taken so far under this Work Package;
- Section 3 provides an overview of the methodology for the case studies that will be used to measure and value elements of the system;
- Section 4 provides an overview of the methodology for the measurement and valuation which is a crucial element related to the wider cost effectiveness framework;
- Section 5 describes the next steps to be followed in this Work Package.

2. Overview of steps taken

The first deliverable of Work Package 14 corresponded to the first objective of the Work Package, namely 'to identify the important elements in calculating costs and benefits of COMPARE and related methods and tools, both regarding the system itself and from the societal perspective'. It was delivered in Month 18, and included a detailed description of the elements of COMPARE and the results of research concerning possible components of the methodological framework for the cost-effectiveness analysis. In parallel to the preparation of Deliverable 1, we also submitted a scientific article on the work conducted so far in WP14, which will be published in the OIE Scientific and Technical Review in 2017.[1]

Since the delivery of Deliverable 1, we have been working on further developing the methodological approach for the cost-effectiveness analysis and on defining the scope and focus of the case studies to measure and value elements of the system which will be conducted during the next project phase of WP 14. Furthermore, a modelling study has been conducted in which costs and benefits of early and targeted in an Ebola outbreak were estimated. Finally, a literature review has been conducted with the aim to make an inventory of methods currently used to estimate the value of safety.

[1] Developing a framework to assess the cost-effectiveness of COMPARE – a global platform for the exchange of sequence-based pathogen data, to be published in OIE Scientific and Technical Review, Vol. 36 (1).

3. Methodology for measuring and valuing elements of the system using case studies[2]

In this section we first discuss the objectives and intended scope of the case studies that will be used to measure and value elements of the system, before describing criteria for the case study selection and providing an overview of the five candidate case studies under consideration. Finally, we describe the methodology for the cost-effectiveness estimation.

3.1 Objectives and scope of the case studies

According to the description of Work Package 14, the cost-effectiveness of COMPARE and related methods and tools will be estimated using case studies, which could be 'retrospective studies (e.g. in relation to past outbreaks such as SARS or H1N1 (swine flu)), scenario studies (e.g. covering specific outbreak scenarios in countries with less developed surveillance systems), and pilot studies, relating to actual uses of the COMPARE system'. In addition, the WP description highlights the specific character of the case studies, emphasising that 'the case studies will focus on a given element of the approach, situation, pathogen, user situation/perspective, or region (developed or developing)'.

For the definition of the scope of the case studies and the methodology for the cost-effectiveness estimation it is essential to clarify the following aspects of the analytical framework (see also Deliverable 1):

- *Definition of the system* to be assessed;
- *Scope of activities* to be assessed;
- *Perspective* of the analysis;
- *Time period covered and geographic scope.*

These aspects are separately discussed in the following sub-sections.

3.1.1 Definition of the system subject to the cost-effectiveness estimation

In Deliverable 1, we have described the elements of COMPARE according the project proposal and listed the related Work Packages and activities. A brief summary is provided in the table below, differentiating between the COMPARE framework (the essential elements including management of the Consortium) and the supporting activities.

[2] Responsible partner for section: Civic Consulting.

Table 1: Overview of elements of COMPARE and related Work Packages and activities

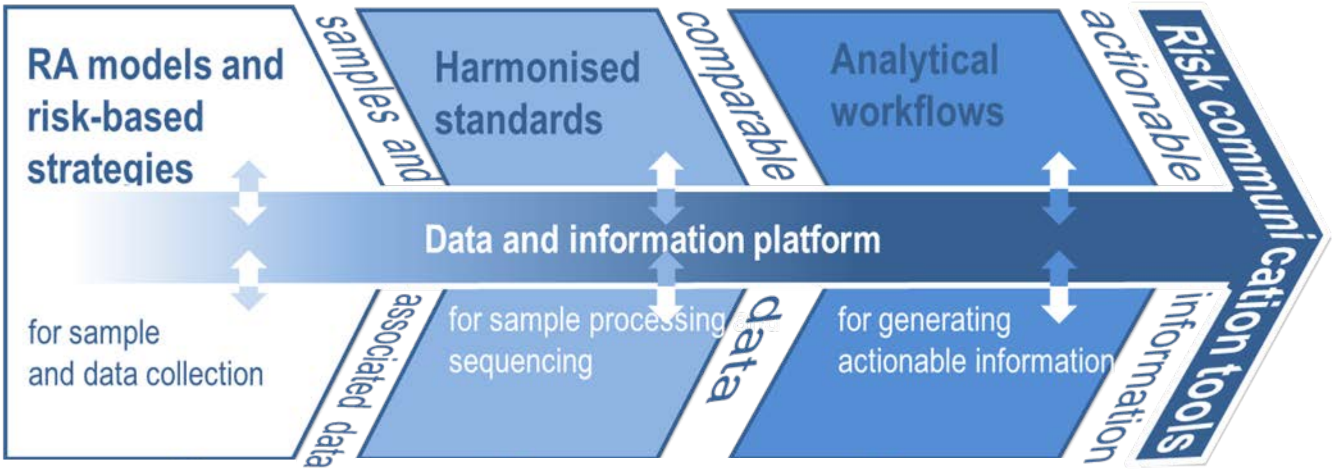
Level	Element	Corresponding Work Package
Compare framework	A. Risk-assessment models and risk-based sample and data collection strategies	WP1 RA and risk-based sample and data collection
	B. Harmonized standards for sample processing and sequencing	WP2 harmonised standards sample processing and sequencing
	C. Analytical tools and methods for sequence-based pathogen and outbreak detection and analyses	WP3 Analytical workflows frontline diagnostics (underpinning research studies in WP6)
		WP4 Analytical workflows foodborne outbreaks (underpinning research studies in WP7)
		WP5 Additional tools for detection of and response to (re-) emerging infections (underpinning research studies in WP8)
	D. Data information sharing platform	WP9 Data and Information sharing platform
	E. Risk communication toolbox	WP10 Risk Communication
	M. Management	WP15 Consortium Management
Supporting activities	F1. Consultations	WP11 User-Stakeholder Consultations
	F2. Research on barriers to open data sharing	WP12 Barriers
	F3 Training and dissemination	WP13 Dissemination & Training
	G. Cost-effectiveness analysis	WP14 Cost-effectiveness

Source: Civic Consulting, adapted from COMPARE proposal. For more detailed information regarding the related work packages and activities, see Deliverable 1 and the COMPARE proposal.

As the table illustrates, COMPARE as defined in the project proposal is complex and includes a number of components that are more related to its character as a research project (e.g. consortium management, consultations and also the cost-effectiveness analysis), than being elements of the core system as such. For the cost-effectiveness analysis, however, a clear definition of the 'system' subject to the assessment is needed.

The following diagram from the COMPARE project proposal illustrates how the core system could be understood:

Figure 1: The core COMPARE system



Source: adapted from COMPARE proposal.

When interpreted in conjunction with the elements of the COMPARE project presented in the previous table, the core system can be understood as a *process of information creation and analysis for pathogen identification and surveillance based on Whole Genome Sequencing (WGS),*[3] starting with risk-based sample and data collection strategies, continuing with sample processing and sequencing based on harmonised standards, and aiming at generating actionable information for pathogen and outbreak detection and related risk communication – all facilitated by a data and information sharing platform. To analyse the cost-effectiveness of this process requires its practical application in specific situations and geographical areas, which is complicated by the fact that the COMPARE project is very broad in scope (being a cross-sector and cross-pathogen framework for a globally linked data and information sharing platform), and does not only concern the practical implementation of such a system, but also intends to first develop the necessary standards and tools for sampling, processing, sequencing and data analysis and interpretation. In line with the case study approach for the cost-effectiveness estimation it was therefore decided to focus on specific application cases with a defined sectoral and geographic scope, considering specifically the costs and benefits of a routine use of WGS for pathogen identification and surveillance. As the routine use of WGS in pathogen identification and surveillance is still relatively rare, it was also decided to widen the perspective to also include other systems that are similar to COMPARE, in that they involve harmonised methods and the use of a centralised data and information sharing platform, even if they are not directly linked to COMPARE (such as the US Genome Trakr network), or have a narrower geographic or sectoral scope (such as purely national systems, see below). In conclusion, it was decided to define the 'system' to be assessed in the cost-effectiveness case studies as follows:

- A system for pathogen identification and surveillance using WGS on a routine basis;
- With harmonised methods (e.g. regarding data collection strategies, sample processing and sequencing, analytical tools and methods); and

[3] The most advanced group of methods for performing whole-genome sequencing (WGS) of pathogens, next-generation sequencing (NGS), enables researchers to generate 'complete genomic information from the isolate or sample independent of both the sector (public health, veterinary health, food safety), and the type of pathogen (viruses, bacteria, parasites). The outputs (sequence data) provide one common language that can be exchanged and compared between laboratories and over time'. See COMPARE project proposal.

- Using a centralised data and information sharing platform for sequences and related metadata ('data hub').

This definition of the 'system' subject to the cost-effectiveness estimation (hereafter referred to as 'WGS-based surveillance system') has been used for the identification of suitable case studies and for refining the methodological approach for the case studies, as elaborated in the subsequent sections.

In discussions within the research team, with COMPARE partners and the COMPARE consortium management, it also became clear that in addition to better understanding the costs and benefits of the practical application of a WGS-based surveillance system, a key research interest would also be to explore potential efficiency gains compared to surveillance systems based on other analytical methods, due to the increased amount of information available through the process (as WGS is expected to facilitate e.g. earlier identification of pathogens and outbreak clusters, a better traceback to outbreak sources, more targeted sampling, etc.), and the possibility that WGS provides to have surveillance systems with a cross-pathogen focus. Taking these considerations and the work package description of WP14 into account, the key research questions for the cost-effectiveness case studies for measuring and valuing elements of the system are therefore defined as follows:

1. What are the costs and benefits of a system for pathogen identification and surveillance using WGS on a routine basis, using harmonised methods and a central data hub, compared to a system that uses other ('traditional') analytical methods?
2. Does the use of WGS lead to cost savings and other benefits due to more targeted sampling and earlier/more effective outbreak response in different contexts (e.g. animal health, food safety, public health), due to the additional information obtained?
3. Are there potential efficiency gains when upgrading established surveillance systems with WGS through converting disease/target population specific surveillance systems into cross-pathogen 'catch all' surveillance systems, including in the context of low and middle income countries?
4. What main factors affect the costs and benefits of a WGS-based surveillance system, and what are possible implications in view of improving its overall cost-effectiveness?

3.1.2 Scope of activities to be assessed

A system for pathogen identification and surveillance using WGS on a routine basis can be conceptualised as consisting of a process of data flow, leading from data acquisition (sampling and sequencing) to data analysis and storage (bioinformatics analysis, data hub) to data application (for outbreak identification and response, but also for other applications). This is depicted in more detail in the table below, which lists the three levels of data flow, and the related key steps and activities. It takes into account the stylised data flow through the COMPARE system as outlined above, but has been further refined based on the literature reviewed and interviews conducted to reflect the data flow in existing WGS-based surveillance systems in generalised terms. For example, an intermediate step of pathogen identification with standard methods is added, because in some surveillance systems samples are first screened with standard methods before positive samples are further analysed with WGS.

Table 2: Process of data flow for WGS-based surveillance systems

Data flow level	Key steps	Includes activities**
1. Data acquisition	Sampling	Risk-based sampling strategies Sample collection Storage and transport Metadata collection
	Pathogen identification*	Primary identification of pathogen with standard methods (e.g. PCR, PFGE, etc.), where relevant Cultivation to obtain isolates for further investigation, where relevant
	Sequencing	Sample processing (incl. DNA extraction and purification) Library preparation Sequencing
2. Data analysis and storage	Bioinformatics analysis	Quality control Data assembly Result validation
	Reference database	Comparing sequence data with reference database Data storage
3. Data application	Outbreak identification and response	Finding clinical links Identifying clusters Conducting tracebacks Containment measures
	Other applications of genomic data	Developing new diagnostic methods and tests Other uses of data (research, etc.)

Source: Civic Consulting, adapted and significantly changed from Allard M. et al 2017. Notes: *Pathogen identification with standards methods may not be relevant in all cases, e.g. when causative agents are unknown. **Not all activities are listed, and some of the listed activities are not relevant in all cases

Most of the key steps and related activities listed in the table above are relevant for any surveillance system using WGS, independent from whether all steps and activities are conducted by one institution (as could be the case for e.g. a national food control laboratory integrated in a public food safety authority), or whether separate institutions are involved in data acquisition, data analysis and storage and data application (which would be the case for example in a network of partner institutions using a common reference database, as is the case in the COMPARE system, or the Genome Trakr network, see below).

With the exception of the steps that are directly related to the Whole Genome Sequencing (such as bioinformatics analysis), other key steps and the related data flow from sampling to outbreak identification and response also characterise surveillance systems that use 'traditional' (non-WGS) analytical methods. In other words, the main differences between the two approaches are at the data analysis and storage levels, although the more granular and detailed information of sequence-based approaches may also influence sampling and response. The model of data flow and the related steps and activities depicted in the table above can therefore be used as analytical basis for identifying the costs and benefits of WGS-based surveillance systems as well as of traditional surveillance systems.

As has been discussed in Deliverable 1, the analysis of costs and benefits of an intervention requires the definition of a baseline (also known as the counterfactual scenario or the comparator). In this analysis the baseline will in most cases be the surveillance system subject to a case study *without WGS, using the next-best method for pathogen identification* (hereafter referred to as a 'traditional surveillance system'). For example, for a case study conducted in an ex-post perspective, e.g. for a national surveillance system where WGS has already been introduced for routine use, the baseline is the system that would be in place if WGS had not been introduced. This could be a system using an analytical method that was in place before introduction of WGS, or which applies a more recent standard method (the above mentioned 'next-best method'). In both cases the aim of the analysis will be to identify changes in costs and benefits for each key step highlighted in the table above between the WGS-based surveillance system and the baseline system. In other words, the focus of our analysis will be on the measurement and valuation of the *marginal (incremental) costs and benefits of introducing WGS* in the surveillance systems subject to our case studies. This has the advantage that costs and benefits that are clearly unaffected (such as costs for depreciation of laboratory buildings) do not need to be assessed, allowing the analysis to focus in detail on those costs and benefits where changes occurred.[4]

A key question in terms of scope is to which extent response activities have to be considered when analysing the costs and benefits of a WGS-based surveillance system (i.e. the degree to which level 3 - data application should and could be included in the assessment). The COMPARE system as depicted in Figure 1 above leads to actionable information for outbreak detection and analysis as well as risk communication, and does not necessarily include other outbreak response measures. The WHO Guide on evaluating the costs and benefits of national surveillance and response systems concludes that an analysis of costs and benefits should consider surveillance and response systems together.[5] While the guide acknowledges that it is possible in theory to evaluate the costs of surveillance and response systems separately, this is not recommended, as the costs of the response system are inextricably linked to the information that is provided by the surveillance system. For the assessment of benefits, it is also not possible in practice to separate the value of the benefits of surveillance from the value of the benefits of response, as the main purpose of the information generated by the surveillance system is to improve the quality of the response. Indeed, the use of WGS influences the way that diagnostics and sample analyses are conducted, and the additional information obtained by WGS regarding the pathogen in turn may influence outbreak identification, the analysis of data and interpretation of trends, and decision-making regarding the response. For example, in the experience of the Genome Trakr network, WGS data was able to provide a clearer distinction between cases and foods that are likely part of a given outbreak and those that are not.[6] On the other hand, data on costs and benefits of response activities are often very difficult to obtain ex-post, and measurement problems are significant, mostly due to the need to assess an appropriate counterfactual (such as the size of an outbreak if a specific response measure had not been taken). In the case studies, we will therefore explore to which extent outbreak response measures can meaningfully be included in the scope of the research, depending e.g. on whether relevant data on costs and benefits of response measures are available, or whether modelling results by WP partner EUR can be applied for this purpose.[7]

3.1.3 Perspective of the analysis

A common perspective adopted in health economics is that of the single benevolent decision-maker, whose aim is to maximise population health based on the available resources. The benefits taken into account are those that accrue to the target population; the costs are those provided for in the health budget. The benevolent decision-maker, in other words, considers the ‘investment case’ for implementing the new system (here: the WGS-based surveillance system). This perspective of the benevolent decision-maker attempting to optimise the allocation of health resources for the benefit of the population will be the approach taken in the case studies.

[4] For case studies in which the surveillance system would not exist without WGS (such is the case for the Genome Trakr network), it is likely that the scenario of a situation without the system is a more appropriate baseline. This will be decided once the initial research on the network is concluded.

[5] World Health Organisation, Evaluating the Costs and Benefits of National Surveillance and Response Systems: Methodologies and Options, 2005, p. 10-1.

[6] Allard M. et al 2017, p. 1978-9.

[7] EUR is conducting complementary modelling of costs and benefits of WGS-based surveillance and response in a societal perspective for selected pathogens (see section 4).

However, following the recommendation of the WHO to adopt a broader 'social perspective' where possible,[8] our approach will take into account the broader benefits for society where this can be feasibly calculated or data from modelling exercises in this respect are available.

3.1.3 Time period covered and geographic scope

As has been highlighted in Deliverable 1, the timeframe chosen to conduct an analysis of costs and benefits may significantly affect results. Evaluating an intervention for a typical (or reference) year may bias the outcome against programmes that take a number of years to start providing benefits to its users. This is particularly the case for surveillance systems, which are intended to be long-lasting, which take time to become effective, and which deal with disease outbreaks that may occur either rarely or on multi-year cycles.[9] On the other hand, a long evaluation period is less feasible in practice.[10]

The WHO recommends selecting a reference timeframe based on the purpose of the study and the availability of data, but suggests that for evaluations of early-warning systems for epidemic-prone diseases, the selected timeframe should ideally include time for at least one outbreak of each major cyclical disease.[11] Given the broad range of institutions and diseases under consideration in the case studies, the determination of the timeframe will be made on a case-by-case basis depending on the specific features of the surveillance system considered and the availability of data. One possible option is to focus the analysis on both the identification of set-up costs (i.e. one-off costs for introducing WGS for routine use in a surveillance system) as well as on recurring costs and benefits of conducting routine surveillance activities with the use of WGS accruing in a typical year. However, as many of the key benefits of WGS are expected to manifest themselves in the identification of outbreaks and the implementation of response measures (especially in terms of averted cases, see below), complementary data on outbreak response, outbreak costs and cases of illness relevant for the case study will be collected and considered, where available, focusing on a suitable period before and after the introduction of WGS for pathogen identification and surveillance.

Once data collection is complete and the costs and benefits have been calculated individually for each case study, the overall results of the case studies will then be compared, possibly by extrapolating the results to a common timeframe, if this turns out to be feasible.

A similar consideration is valid for the geographical scope to be chosen, which will also depend on the case studies finally selected. The assessment will focus on the most suitable geographical coverage, which will often be in line with the coverage of the surveillance system subject to the case study, or a suitable part of it (likely to be a specified country or region). Where relevant spillover effects occur, i.e. costs and benefits that accrue beyond the 'borders' of the particular surveillance system, they will be considered to the extent that data is available in this respect.

---

[8] Edejer T. Tan-Torres, R. Baltussen, T. Adam, R. Hutubessy, A. Acharya, D.B. Evans, and CJL. Murray, Making Choices in Health: WHO Guide to Cost-Effectiveness Analysis, 2003, p. 18-9.

[9] World Health Organisation, Evaluating the Costs and Benefits of National Surveillance and Response Systems: Methodologies and Options, 2005, p. 18.

[10] See also: Edejer T. Tan-Torres, R. Baltussen, T. Adam, R. Hutubessy, A. Acharya, D.B. Evans, and CJL. Murray, Making Choices in Health: WHO Guide to Cost-Effectiveness Analysis, 2003.

[11] World Health Organisation, Evaluating the Costs and Benefits of National Surveillance and Response Systems: Methodologies and Options, 2005, p. 18.

3.2 Selection of case studies

The following section describes possible candidates for the cost-effectiveness case studies for measuring and valuing elements of the system. The section first sets out the criteria used for the identification of these case studies and provides then the profiles of the candidate case studies currently under consideration.

3.2.1 Criteria for case study selection

In line with the definition of the system and the scope of activities to be assessed in the cost-effectiveness estimation as described in the previous section, and to cover a diversity of situations, the criteria for case study selection were as follows:

- *Type of surveillance system.* The candidate case studies should represent surveillance systems (i.e. with a distributed network of members) or reference centres (without distributed network) that use WGS on a routine basis, with harmonised methods and a central data hub;
- *Sector of application*. The candidate case studies should represent WGS-based surveillance systems or reference centres in different sectors which cover pathogens that are relevant in a public health, food safety, and animal health perspective;
- *Geographical level*. The candidate case studies should cover WGS-based surveillance systems or reference centres at the national, regional (EU), and international level.

Based on these criteria, a total of five candidate case studies were identified (see next sub-section), which are focusing on Salmonella and Avian Influenza.[12] Both pathogens were identified as relevant pathogens in terms of public health, food safety and animal health/animal production importance, as well as pathogens where WGS-based surveillance systems or reference centres already exist. From these five candidate cases studies, a final set of two to three case studies will be selected after the conclusion of the screening process, also depending on the willingness of the institutions hosting/coordinating the identified surveillance systems to participate in a case study. Note that the current selection of case studies does not include a case study in a low and middle income country, as no obvious candidate for a system or reference centre using WGS on a routine basis could be identified so far. Research in this respect is, however, ongoing.

3.2.2 Description of candidate case studies

The table below presents the candidate case studies that were identified and are currently being screened. It provides information on the organisation, the country and geographical level, as well as the pathogen focus of the case study. The assessment of costs and benefits of WGS-based surveillance will focus on this pathogen (as indicated above, either Salmonella or Avian Influenza), while possible efficiency gains due to a cross-pathogen focus will be also explored, where more than one pathogen is covered.

[12] *Salmonella* bacteria are the most frequently reported cause of foodborne outbreaks with known origin in the EU. Each year, more than 100,000 human cases are reported in the EU, with the overall annual economic burden estimated at up to EUR 3 billion. *Avian influenz*a is a highly contagious viral infection affecting wild and domestic birds. Due to the related economic burden and its zoonotic potential, Avian influenza is subject to strict mandatory surveillance and reporting within the EU.

Table 3. Overview of the candidate case studies

Organisation	Country	Geographical level	Pathogen selected for case study	Cross-pathogen focus	Status
Danish Veterinary and Food Administration*	Denmark	National	Salmonella	Yes	Contacted, declared willingness to participate in case study
Public Health England - Salmonella surveillance network	UK	National	Salmonella	Unclear	In process of being contacted
FDA - Genome Trakr network	USA	National/ international	Salmonella	Yes	Contacted, ongoing discussions concerning possible case study participation
Friedrich Loeffer Institut**	Germany	National	Avian Influenza	Yes	Contacted, ongoing discussions concerning possible case study participation
The Animal and Plant Health Agency** - EU-RL network for avian influenza	UK	UK/EU	Avian influenza	No	Contacted, ongoing discussions concerning possible case study participation

Source: Civic Consulting. Notes: *Danish Veterinary and Food Administration cooperates with COMPARE partner DTU for analysis of sequencing data and therefore is a user of the COMPARE system. **FLI and APHA are both COMPARE partners.

As indicated in the last column of the table, the host organisations of all but one of the candidate case studies have been contacted and are either in the process of discussing their possible participation, or have already declared their willingness (in principle) to be subject to a case study.

Each of the listed candidate case studies is briefly discussed in the following pages.

Danish surveillance programme for salmonella

Salmonella causes the second highest number of cases of foodborne illness in Denmark (the first being Campylobacter). Since the early 1990s, action plans have been in place to control salmonella, and plans have been established in the poultry, pig and cattle production. These action plans have been successful in reducing the number of human cases.[13]

The microbiological laboratory of the Danish Veterinary and Food Administration of the Ministry of Environment is in charge of conducting official controls in Denmark, performed to ensure the verification of compliance with feed and food law, animal health and animal welfare rules. While certain laboratory activities related to these controls used to be outsourced to the Technical University of Denmark (DTU), the Food Administration and its own microbiological laboratory recently initiated a process of transferring related activities back in-house. In the context of this reallocation of activities, it was decided to invest in WGS for microbiological analysis, following

---

[13] Ministry of Environment and Food of Denmark, https://www.foedevarestyrelsen.dk/english/Food/Bacteria-and-vira/Pages/default.aspx (last accessed on 12.06.2017)

observations made by EFSA and the WHO and with the aim to develop broader methods that would allow for more and better information to be obtained from the analysis of samples. From January 1, 2017, MLVA is no longer used for molecular typing of Salmonella from food, feed and animal samples. Instead, isolates are analysed by WGS. Primary analyses of samples and sequencing of isolates are performed by the microbiology lab of the Food Administration and the bioinformatics functions are hosted at the DTU.[14]

Public Health England Salmonella surveillance network

The UK's national reference laboratory (NRL) for food microbiology is responsible for official controls on specific areas of food microbiology: Listeria monocytogenes, coagulase positive staphylococci, Escherichia coli (including STEC), campylobacter, salmonella, and antimicrobial resistance. The Food Standards Agency (FSA) awarded PHE these responsibilities for performing official controls to ensure the verification of compliance with feed and food law, animal health and animal welfare rules, as described in Article 33 of Regulation (EC) No 882/2004.

In 2012, Public Health England developed a central genomics service within the reference microbiology laboratories at PHE Colindale. The focus was placed on a few pathogens, including Salmonella spp. The central genomics service was launched in April 2014 and sought to validate their processes through running a pilot *Salmonella* typing project.[15] WGS has now been implemented for routine identification, characterisation and typing of Salmonella isolates.[16] Salmonella isolates from England, Wales, and Northern Ireland are served by the Gastrointestinal Bacteria Reference Laboratory in Public Health England, where whole genome sequencing is performed, and MLVA has been discontinued. The Scottish Microbiology Reference Laboratories also intend to implement WGS for Salmonella in 2017.[17]

Genome Trakr network

The GenomeTrakr network, created by the US FDA, is a distributed network of laboratories that utilises Whole Genome Sequencing for pathogen identification in foodborne outbreaks. The network consists of public health and university laboratories that collect and share genomic and geographic data from foodborne pathogens. As of April 2017, the network consists of 15 federal labs, 25 state health and university labs, 1 U.S. hospital lab, 2 other labs located in the U.S., 20 labs located outside of the U.S., and collaborates with independent academic researchers.[18] Data curation and bioinformatic analyses and support are provided by the National Center for Biotechnology Information (NCBI) at the National Institutes of Health. The FDA Foods Whole Genome Sequencing Staff coordinates efforts to sequence pathogens collected from foodborne outbreaks, contaminated food products and environmental sources. The genome sequences are archived in the GenomeTrakr open-access genomic reference database, which can be used for real time comparison and analysis.

As of today, the GenomeTrakr network has sequenced more than 113,000 isolates and is regularly sequencing over 3,500 isolates each month. The current GenomeTrakr database contains roughly 33,000 Salmonella

---

[14] EURL-Salmonella Newsletter March 2017 and exploratory interview Head of Unit of the Microbiological laboratory, Danish Veterinary and Food Administration.

[15] Pathogen Genomics Into Practice, PHG Foundation, 2015.

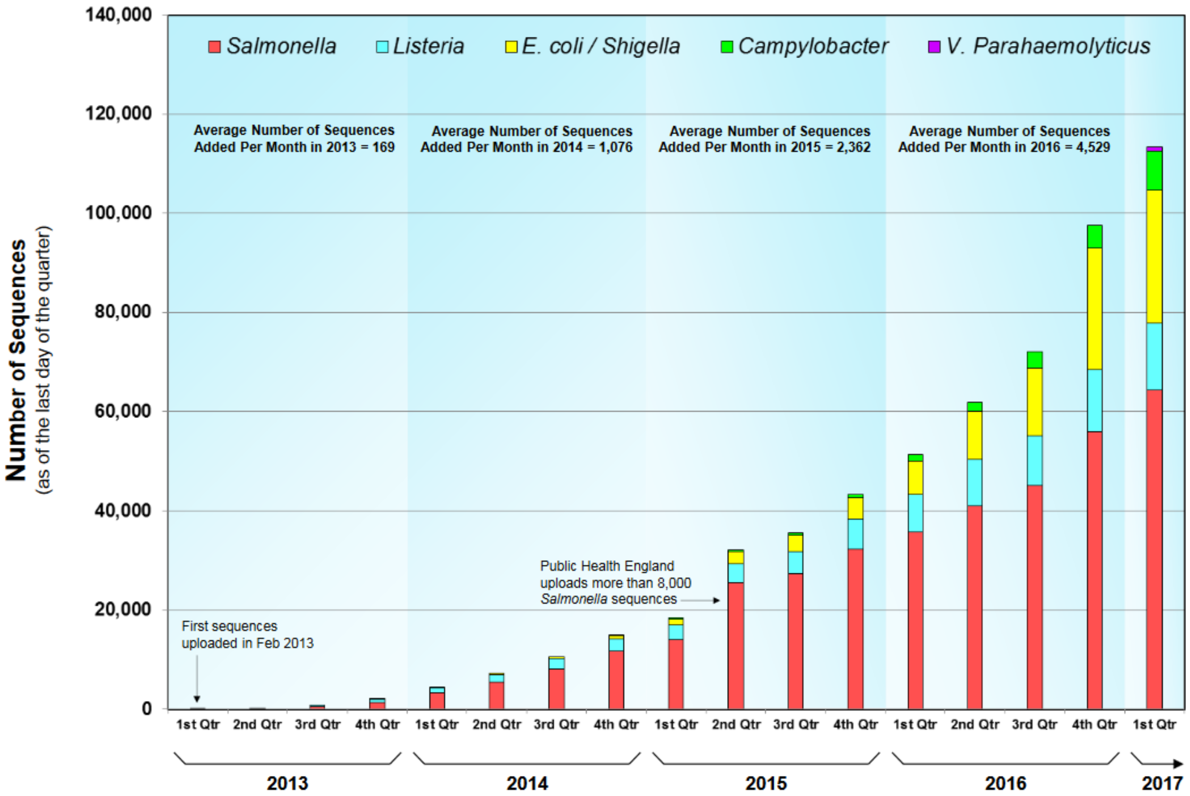
[16] Ashton, P., Nair, S., Peters, T., Tewolde, R., Day, M., Doumith, M., Green, J., Jenkins, C., Underwood, A., Arnold, C. and de Pinna, E., 2015. Revolutionising public health reference microbiology using whole genome sequencing: Salmonella as an exemplar.

[17] EURL-Salmonella Newsletter March 2017

[18] U.S. Food and Drug Administration,
https://www.fda.gov/food/foodscienceresearch/wholegenomesequencingprogramwgs/ucm363134.htm

isolates, 7,000 Listeria isolates, 5,000 E. coli/Shigella isolates, and 1,000 Campylobacter isolates.[19] The figure below provides an overview of the number of sequences in the database over time.

Figure 2: Total number of sequences in the GenomeTrakr Database



Source: U.S. Food and Drug Administration website consulted in May 2017, https://www.fda.gov/Food/FoodScienceResearch/WholeGenomeSequencingProgramWGS/ucm403550.htm

As indicated in the figure above, the project was originally focused on Salmonella and Listeria. The FDA's priority is to expand the network capacity in foods and to equip more state health and agriculture laboratories in order to sequence and include collected pathogens from food and the environment as part of inspection activities, also possibly beyond foodborne pathogens.

Friedrich-Loeffer-Institut

The Friedrich-Loeffer-Institut (FLI) is the National Institute for Animal Health in Germany. It is a federal research institute and independent higher federal authority under the Federal Ministry for Food and Agriculture. Its work aims at the prevention of diseases, the improvement of animal welfare and the production of high quality animal-based foodstuffs. The institute performs epidemiological investigations during outbreaks of animal diseases. It also prepares risk assessments on various infectious diseases of farm animals.

[19] U.S. Food and Drug Administration, https://www.fda.gov/food/foodscienceresearch/wholegenomesequencingprogramwgs/ucm363134.htm

The National Reference Laboratory for AI within the FLI is the direct contact and reference centre for federal and state authorities, especially for questions relating to diagnostics, and is also active within the EU-RL network for Avian Influenza (see below). As a reference laboratory of the World Organisation for Animal Health (OIE) and of the Food and Agriculture Organization (FAO) of the United Nations, the laboratory also provides advice and diagnostic assistance to countries outside Europe. The laboratory conducts application-oriented research in the field of avian influenza virus diagnostics, epidemiology and pathogenesis. Furthermore, the NRL deals with scientific questions relating to porcine influenza.

FLI has a laboratory for NGS and Microarray Diagnostics. The main task of the laboratory for NGS and microarray diagnostics is full-length DNA or RNA virus genome sequencing. Beyond the sequencing activities, establishing new technical equipment, molecular biological methods, and implementing new ways for data analyses are among FLI's focus areas.[20]

The Animal and Plant Health Agency (APHA)

The Animal and Plant Health Agency (APHA) is an executive agency of the Department for Environment, Food & Rural Affairs, and also works on behalf of the Scottish Government and Welsh Government. APHA is responsible for identifying and controlling endemic and exotic diseases and pests in animals, plants and bees, and for surveillance of new and emerging pests and diseases. It conducts scientific research and acts as an international reference laboratory for many farm animal diseases, including avian influenza. APHA is also the EU reference laboratory for avian influenza, and uses WGS in this context. Cases of avian influenza found in the EU are confirmed at the APHA and Member States must report the results of their surveillance to the Commission. The surveillance is compulsory according to Council Directive 2005/94/EC on EU measures for the control of avian influenza that must follow harmonised guidelines laid down in Commission Decision 2010/367/EC.

3.3 Assessment of costs and benefits

The following section presents the specific cost and benefit types to be assessed in the cost-effectiveness case studies for measuring and valuing elements of the system.

3.3.1 Evaluation of costs

As discussed in Deliverable 1, the literature review identified a range of approaches for cost categorisation used in relevant analyses. The WHO Guide to Cost-Effectiveness Analysis, for example, classifies costs according to the resources used (labour, capital, consumables, and overhead costs).[21] In the context of evaluating genomic technologies, however, Buchanan et al. (2013) instead propose cost categories that correspond to the steps involved in medical testing using whole genome sequencing, for example, costs related to sample collection, sample testing, data analysis, communication of text results, etc.[22] For the present assessment, we will use a combination of the two classification approaches that categorises costs according to both the type of resources

---

[20] Friedrich Loeffer Institut, https://www.fli.de/en

[21] Edejer T. Tan-Torres, R. Baltussen, T. Adam, R. Hutubessy, A. Acharya, D.B. Evans, and CJL. Murray, Making Choice in Health: WHO Guide to Cost-Effectiveness Analysis, 2003.

[22] Buchanan, James, Sarah Wordsworth, and Anna Schuh, "Issues Surrounding the Health Economic Evaluation of Genomic Technologies", Pharmacogenomics, Vol. 14, No. 15, 2013, Appendix 3: Costs which could be included in economic evaluations of genomic technologies.
http://www.futuremedicine.com/doi/abs/10.2217/pgs.13.183\nhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3909837&tool=pmcentrez&rendertype=abstract.

used and the key steps in the analysis, focusing on the costs of the system itself (i.e. costs accruing to the institutions involved in the implementation of the surveillance system).[23]

For the classification of costs based on the type of resources used, the following five cost categories have been selected for the assessment based on the WHO guidance above, the literature review conducted for Deliverable 1, and the previous experience of Civic Consulting in analyses of costs and benefits of reference laboratory networks, as well as of veterinary systems:[24]

- *Staff costs.* Staff costs include wages, social contributions and non-wage income of employees, such as in-kind payment;[25]
- *Capital equipment costs*. The reduction in the value of fixed assets includes the depreciation of sequencers and ancillary equipment, IT infrastructure for bioinformatics applications and other fixed assets needed for using WGS on a routine basis (including software, if not purchased as a service);[26]
- *Consumable materials*. Consumable materials are items that are used up in the provision of a good or service. In the context of the case studies, this generally relates to consumables/reagents used in laboratory processes, including test kits, petri dishes, etc.;
- *Sub-contracting and services.* All costs and fees arising from subcontracting a certain activity or parts of a certain activity to private companies or laboratories outside of the institutions that are part of the surveillance system under review are considered. This would include sub-contracting of analytical services, external training courses, external quality assessments by private providers, etc.;
- *Overhead/other costs*. Other cost items which do not fall under one of the other five categories, including travel costs, shipping costs, interests, utilities, rents, maintenance, etc.[27]

The classification of costs based on key steps of the process which constitutes the surveillance system follows the process of data flow depicted in Table 2 above, i.e. differentiating between:

- *Data acquisition.* Sampling, primary pathogen identification, sequencing;
- *Data analysis and storage.* Bioinformatics analysis, reference database;
- *Data application.* Outbreak identification and response, other applications of genomic data.

Specific cost data for each case study will be collected from the relevant institution(s). We will clarify in dialogue with the case study institutions in which format they could provide data, and to which extent they could provide

[23] Cost (savings) accruing to the wider society due to more targeted surveillance and potentially reduced outbreak costs are considered in the context of assessing benefits.

[24] Civic Consulting (2016), Study on cost–benefit analysis of reference laboratories for human pathogens: final report, study conducted for CHAFEA of the European Commission and Civic Consulting (2009), Cost of National Prevention Systems for Animal Diseases and Zoonoses in Developing and Transition Countries, study conducted for the OIE.

[25] If no specific data on staff costs in monetary terms is available, staff costs will be estimated on basis of staff category and time, using standard labour cost data.

[26] For the calculation of capital equipment costs, the initial, one-off costs that are incurred when introducing WGS in a surveillance system are calculated and the reduction in value is calculated either based on data provided by the institution, or calculated based on the average service life of the equipment (e.g. 3 to 8 years, depending on the type of equipment).

[27] Specific cost items, such as overheads for use of office space and standard office equipment, will be estimated if no specific data is available.

data on investment costs for introducing WGS and possibly multi-annual data on recurring costs. Where needed, we will convert the data provided into a standard format to ensure consistency between the case studies.

Cost data per key step and per type of resource will be collected from the case study institutions with the help of a dedicated questionnaire which is based on a cost matrix, and which combines the categorisation of costs according to the type of resources used and the key steps in the analysis (see Table 4 on the following page). This will allow us to draw conclusions regarding the main resource types that determine the additional costs of WGS-based surveillance systems, as well as identifying the key cost items in terms of main steps/activities conducted. Note that the table presented on the following page is an indicative version of the cost matrix, which will be further refined during a pilot case study.

Table 4. Indicative cost matrix for case study data collection regarding costs

Key steps and related activities			Cost type by resource						*TOTAL cost by key step*	Comments
			Relevant?	Staff costs	Equipment costs	Consumables	Subcontrac-ting/ services	Overheads/ other costs		
1. Data acquisition	Sampling	Risk-based sampling strategies Sample collection Storage and transport Metadata collection	Y/N							
	Primary pathogen identification	Use of standard methods Cultivation to obtain isolates	Y/N							
	Sequencing	Sample processing Library preparation Sequencing	Y/N							
2. Data analysis and storage	Bioinformat-ics analysis	Quality control Data assembly Result validation	Y/N							
	Reference database	Comparing sequence data with database Data storage	Y/N							
3. Data application	Outbreak identification and response	Finding clinical links Identifying clusters Conducting tracebacks Containment measures	Y/N							
	Other appli-cations of genomic data	Developing new diagnos-tic methods and tests Other uses of data (re-search, etc.)	Y/N							
TOTAL cost by resource type										

Source: Civic Consulting. Notes: *Only incremental costs resulting from the introduction of WGS are considered in the analysis of cost.

3.3.2 Evaluation of benefits

The benefits of interventions such as pathogen identification and surveillance systems and/or the introduction of new genomic sequencing technologies are far-reaching and numerous; quantifying such benefits is therefore a significant challenge. The WHO Guide on evaluating the costs and benefits of national surveillance and response systems, for example, lists the following types of benefits that may arise as a result of such systems:

- Benefits derived from *averting cases of illness*;
- Benefits derived from *averting deaths*;
- Benefits of *fewer social and economic disruptions* (including disruptions to trade and tourism) when epidemics are averted; and
- *Social and psychological benefits* stemming from less apprehension and greater peace of mind when large outbreaks of serious infectious diseases are rare or non-existent.[28]

The list above reflects the broader benefits that individuals, institutions, and societies enjoy when infections or outbreaks are averted, including nonmarket goods and services (such as the feeling of safety) which cannot be readily estimated using market prices and budgets. From the point of view of laboratories and public health authorities, however, there may also be potential institutional benefits related to efficiency gains, such as cost savings/avoidance, time saved, and increased productivity.[29] The cost-benefit analyses evaluated as part of the literature review conducted for Deliverable 1 tended to quantify and monetise certain key benefits, describing additional benefits qualitatively.

In line with the perspective chosen for the analysis as discussed above, the present assessment will evaluate benefits falling within the following three categories:

- Benefits resulting from institutional *cost savings/efficiency gains*;
- Benefits resulting from *cases of illness averted*;
- *Other benefits* identified during the course of the case studies.

The three benefit types above accrue to different actors and pose varying degrees of difficulty related to data collection, quantification and monetisation. The subsections below describe each of the benefit types and the associated data collection process.

Cost savings and efficiency gains

Benefits resulting from institutional cost savings and efficiency gains accrue principally to the institutions charged with pathogen surveillance and outbreak response. Examples of potential benefits that fall into this category include:

[28] World Health Organisation, Evaluating the Costs and Benefits of National Surveillance and Response Systems: Methodologies and Options, 2005.

[29] Cellini, Stephanie Riegg, and James Edwin Kee, "Cost-Effectiveness and Cost-Benefit Analysis", Handbook of Practical Program Evaluation, 2010, pp. 493–530.

- *Reductions in per sample testing costs*, for example, through reduced staff and processing time required to prepare and test a sample, and the possibility to perform cross-pathogen analyses;
- *Fewer additional tests* needed to analyse the characteristics of a particular pathogen isolate (e.g. reduced need for complementary tests regarding antimicrobial resistance);
- *More efficient sampling* through the ability to better target higher-risk areas, companies, production processes, animals, persons, or foods;
- *More rapid and targeted outbreak response* as a result of being able to better identify the source of infection or contamination.

Cases of illness averted

Benefits resulting from cases or infections averted accrue to the population more generally in the form of the avoided costs of an infection or outbreak. Examples of benefits (avoided costs) that fall into this category include:

- *Avoided medical costs*, i.e. for diagnosis, hospitalization/treatment costs, medication, etc.;
- *Avoided productivity losses* for the broader economy;
- *Avoided destruction* of infected animals or food products;
- *Avoided long-term human health complications* and premature deaths.

Other benefits

The interviews conducted so far indicated that other benefits related to the use of WGS in a surveillance system may arise related to applications of genomic data for the development of new diagnostic methods and tests, development of new vaccines, or for other research purposes.

Data on benefits will be collected from the relevant institutions for each of the selected case studies, complemented by statistical data available from central institutions, e.g. regarding labour costs and the burden of disease/outbreaks.

3.3.3 Cost-effectiveness estimation

Based on the data collected, we will first separately assess the costs and the benefits of using WGS on a routine basis for pathogen identification and surveillance for each case study and in a second step discuss on this basis the cost-effectiveness of WGS-based surveillance systems. In health economics, units of effectiveness often relate to improvements in health or in the quality of life. Commonly used units in the area of human health include metrics such as disability-adjusted life years (DALYs) and quality-adjusted life years (QALYs).[30] However, in evaluations of interventions relating to animal health, often economic valuation techniques are used to quantify benefits in monetary terms.[31] As the case studies aim at reflecting a cross-sectoral and broad social perspective, the assessment will take into account diverse types of benefits including both health and non-health outcomes. For these reasons, relevant benefits in

[30] Edejer T. Tan-Torres, R. Baltussen, T. Adam, R. Hutubessy, A. Acharya, D.B. Evans, and CJL. Murray, Making Choice in Health: WHO Guide to Cost-Effectiveness Analysis, 2003. p 50.

[31] Babo Martins, Sara, and Jonathan Rushton, "Cost-Effectiveness Analysis: Adding Value to Assessment of Animal Health, Welfare and Production", Revue Scientifique Et Technique-Office International Des Epizooties, Vol. 33, No. 2201, 2014, pp. 1–18.

our analysis will be quantified and monetized, where possible. However, where information on benefits is not readily quantifiable, benefits will be assessed in qualitative terms.

It is expected that a significant part of the benefits resulting from the use of WGS for pathogen identification and surveillance will be in the form of avoided costs from cases of illness averted. The calculation of related benefits necessarily relies on a comparison against a hypothetical counterfactual, which can be challenging for outbreak situations involving complex interactions of outbreak causes, situations and response measures taken. For this reason, the primary means of presenting the cost-effectiveness of the system will be in the form of a break-even analysis. This analysis compares the costs of a surveillance system using WGS (based on the data collected during the case study) with the known costs of infection (e.g. by using standard cost data regarding costs of illnesses) to calculate the break-even point, i.e. the minimum number of cases (infections) that would need to be averted in order for the use of WGS to be cost-neutral. If the number of cases averted can be reasonably expected to exceed this break-even point, then a WGS-based surveillance system is considered to be cost-effective under the particular circumstances of the case study. This break-even analysis will be complemented by reference to the results of modelling of wider societal costs and benefits provided by WP partner EUR, where available.

4. Measurement and valuation of elements related to the wider cost effectiveness framework[32]

4.1 Value of safety

Many of today's societies are governed by rules, regulations and protocols, many of which are designed with the aim of keeping citizens 'safe'. In the dictionary, safety is defined as 'the condition of being protected from or unlikely to cause danger, risk or injury' (Merriam Webster Dictionary 2016). With recurrent news about threats of global warming, terrorist attacks, pandemics and natural disasters, it is no surprise that safety is a significant concern for citizens, companies, and governments. All wish to minimize the possibility of death, illness or injury. However, a question that is increasingly relevant in these same societies is whether policies that aim to increase the safety of citizens, be it in the health, transport or environmental sector, provide good value for money. After all, as public money can be spent only once and investments in increased safety displace investments in other areas of government that may be in similar high demand. In order to evaluate the efficiency of these policies, safety needs to be valued. Due to safety being an intangible, non-monetary good, economists tend to value risk- or uncertainty-reduction instead of 'safety' (Boithias et al. 2016, Henson. 1996, Herzog Jr and Schlottmann. 1990), with risk-reduction being the most tangible and therefore the most applied option in the literature. This being said, there is no 'golden standard' for safety valuation. Early approaches were based on life insurance premiums, which were then replaced, initially by human capital methods, and more recently by stated preference methods (Ball. 2000). This ongoing shift in approaches shows that valuing safety is a complex task.

Inspection of the literature on valuing safety also shows that the research into this topic is scarce, scattered across scientific fields, and that no review of safety valuation literature is currently available. Therefore, the aim of this paper is to present a review of the existent literature; synthesizing the methodologies used in empirical research papers that value safety. The reviewed papers come from different scientific fields, including environmental economics, transport economics, food safety, crime, and health economics; indicating that the results presented in this paper may be beneficial to any future policy or project that requires safety valuation. As the outcomes from these various fields are highly incomparable, the focus of this study is on the methodology of valuation and the characteristics of respondents, context and study design associated with elicited values of safety, as these are the most comparable aspects of all the papers.

While the main aim of this section is to review the methods used in *empirical* research on safety, it is also of value to those developing a value of safety measure to consider some of the main issues brought up in the *theoretical* research in this area. Therefore, the next section of this paper discusses the theoretical background to the valuation of safety. Next, the methods of the literature search are discussed, followed by the findings of the research and finally a discussion section.

4.1.1 Theoretical Background

One of the ways to compare alternative policies or interventions is by using a cost-benefit analysis (CBA) approach, in which the costs and benefits of the alternatives in question are compared between and within alternatives (Cellini and Kee. 2010). In order to compare the benefits from interventions that differ outcome – for example an improvement in road safety versus an improvement in city air quality - we can

[32] Responsible partner for section: EUR.

place these benefits in monetary terms. Doing this for non-monetary goods such as safety has become more important as more public projects are developed and require evaluation (Alpizar et al. 2003). The two main methods of assigning monetary value to non-market goods are revealed and stated preference. Revealed preference uses observed prices and choices to extract the value of a given outcome, while stated preference uses carefully worded surveys to extract how an individual values the chosen non-market good (Brown. 2003). Stated preference is the most common of the two for valuing non-market goods, as it is hard to find observations that are definitively measuring a non-market good such as safety. The most common types of stated preference study used to value non-market goods are contingent valuation (CV) methods and discrete choice experiments (DCE). CV methods directly ask individuals their willingness to pay (WTP) for some non-market good, given a certain hypothetical scenario (Mitchell and Carson. 1989), whereas DCE uses a hypothetical scenario, but asks respondents to choose between options with several different attributes in order to indirectly extract their WTP (Ryan et al. 2007). In the instance of safety valuation, 'safety' is very complex to define, and therefore most papers in this review value a reduction of risk, uncertainty or a specific unsafe incident. However, even with a more tangible definition of safety, several issues still arise when trying to valuate it.

When developing a safety valuation measure it is also important to consider the impact that the design of the study could have on the results. One design feature that has been found to be relevant in safety valuation is the information provided in the survey. Having a clear and comprehensible survey is important, and including too much or too little information about the safety being valued could make questions harder for respondents to understand. Then there is the issue of how to present the information. It can be presented orally, visually or written. Mattea et al. (2016) explore the use of visual information in a stated preference study and find a positive treatment effect linked to the use of visual information when studying risk reduction valuation in landslide programmes (Mattea et al. 2016). Another type of information that is frequently thought of as causing bias in CV results is public opinion. Critics have contested the assumption underlying CV that respondents have 'well-defined and self-interested preferences' and argue that respondents are in fact influenced by public opinion. Chanel et al. (2006) tested this by giving a group of respondents the option to revise their WTP answers after hearing the mean WTP response from the survey group they are in (Chanel et al. 2006). They found that public opinion had no significant impact on respondents' answers. They suggest that it may be a poorly-defined private value structure (or preferences) that leads to a reaction to public opinion (Chanel et al. 2006), which is worth bearing in mind for those developing a CV study.

In CV studies there are also effects to consider such as ordering effects, embedding effects and internal consistency (Halvorsen. 1996). Ordering effects refer to 'where a value of a particular good as perceived by respondents depends on where in the sequence it is valued' (Kahneman and Knetsch. 1992). Embedding effects occur when 'the value of a particular good as perceived by the respondents is sensitive to the number of goods to be valued' (Kahneman and Knetsch. 1992). Internal consistency refers to when the sums of a valuation sequence equal the direct total value. Internal consistency is not frequently tested in CV research, which has worried critics. Halvorsen (1996) researched ordering effects and internal consistency, when testing WTP for 'reduced health damage from air pollution', and found considerable and significant ordering effects which were internally consistent (Halvorsen. 1996). She did not specifically research embedding effects, however she suggested that informing respondents about all effects of a project will reduce embedding effects and that a 'one-short' holistic valuation of all goods may help to mitigate ordering effects.

The way in which an individual values safety is dependent on characteristics of the individual, the context and the study design. The most frequently researched of these characteristics is *risk perception.* This refers to how an individual perceives the level of risk in a situation due to their intellectual judgment (Sjöberg. 1998). High risk-perception (i.e. assuming larger levels of risk than there objectively are) has been shown to lead people to value safety (or risk reduction) more highly (Haddak et al. 2016). It is also worth noting that risk perception is rarely equivalent to worry, as worry is based on emotion rather than intellectual judgment. As Sjoberg (1998) puts it: 'one can feel worried about a risk without believing that it is especially large, and *vice versa*'. However, worry and also pessimism have been shown to be small explanatory factors of risk perception that vary in size depending on the risk being studied (Sjöberg. 1998).

An issue related to risk perception is probability weighting. Individuals are known to not evaluate probabilities linearly (Bleichrodt and Eeckhoudt. 2006) but to overestimate small probabilities and underestimate large probabilities. In fact, Bleichrodt and Eeckhoudt (2006) show that introducing probability weighting strongly affects the WTP for reductions in health risks. Another individual issue to consider is respondent uncertainty. It has been shown that respondents are frequently uncertain about their preferences when answering contingent valuation questions and it is a concern that this uncertainty may be affecting CV results (Logar and van den Bergh, Jeroen CJM. 2012). However, Logar and van den Bergh (2012) found that incorporating information on respondent uncertainty into the model does not lead to any gains compared to a standard CV model.

There are also certain contextual factors that have been shown to influence how individuals value safety. For example, individuals may 'dread' certain situations more than they dread others. It has been argued that in some circumstances, a 'bad death' premium would be relevant when valuing the avoidance of certain types of deaths, such as murder and drowning (Chilton et al. 2006). However, Chilton et al. (2006) found that while dread elements are present in certain types of death, they are often canceled out by low baseline risks. It is therefore worth considering whether to include a 'dread premium' in a study on a case by case basis. Another type of situation worth noting is a *catastrophe.* This refers to occasions when 'large concentrated losses are over-counted relative to dispersed losses' (Zeckhauser. 1996) (consider a plane crash in comparison to a number of car accidents in a year). This means that risk reduction cannot be simply described in terms of reduction of victims, as people who are not directly involved may also consume the catastrophe. For example, there are losses incurred by those who feel empathy for the victims, those who lose financially and those who feel anxiety and worry about the event before it even happens. While this topic is somewhat less directly related to safety, catastrophes clearly have far-reaching spill-over effects and therefore studies valuing reduction in risk of an outcome that may be perceived as a catastrophe may need to include additional information or measures based on the theory discussed (Zeckhauser. 1996).

As previously stated, there is no 'golden standard' for safety valuation. What this shows is that there is clearly more work to be done on methodologies for the valuation of safety, and that current CV research should address the issues discussed in this section to work towards something more closely resembling a 'golden standard' for safety valuation.

4.1.2 Methods

On October 10th and 11th, 2016 a comprehensive literature search for papers related to the valuation of safety was performed using two electronic databases: Embase and Scopus. There was no restriction on time period or language. Book chapters, dissertations, and theses were not considered. In the search the following terms were used: Value, valuation, review, shadow price, willingness to pay, willingness to accept, discrete choice experiment, stated preference, revealed preference, and contingent valuation method. The above terms were used in combination with these search terms: Safety, security, uncertainty reduction, risk reduction. Overall, 40 different combinations of search terms were used to find the papers for this review. Secondary references were found by searching the references of the already included papers in order to find relevant papers that the databases may not have included.

Papers retrieved from the search were selected for review if they fitted both of the following inclusion criteria: (1) The research is empirical, and (2) the research deals with the valuation of safety, security, risk reduction, uncertainty reduction or reduction of some event that is stated to decrease safety. Papers were excluded if safety valuation was not a main objective of the paper or if the paper was not in English (Table 5).

Table 5 Results of Search Terms

Embase	Safety	Security	Uncertainty Reduction	Risk Reduction	Total
Value	29099	2409	15	3312	34835
Valuation	173	61	1	84	319
Shadow Price	1	2	0	0	3
Review	177856	9016	15	24150	211037
WTP	252	24	0	141	417
WTA	41	4	0	8	53
DCE	61	1	0	25	87
Stated Preference	32	1	0	21	54
Revealed Preference	2	0	0	3	5
CV	10	5	0	10	25
Total (incl. Value & Review)					246835
Total (excl. Value & Review)					963

Scopus	Safety	Security	Uncertainty Reduction	Risk Reduction	Total
Value	82152	30435	4535	25783	142905
Valuation	706	1218	143	531	2598

Shadow Price	11	41	4	8	64
Review	194236	20204	1990	67514	283944
WTP	632	181	97	497	1407
WTA	135	58	5	59	257
DCE	93	16	4	70	183
Stated Preference	274	82	13	138	507
Revealed Preference	310	128	8	101	547
CV	85	37	11	87	220
Total (incl. Value & Review)					432632
Total (excl. Value & Review)					5783

One of the authors (MP) screened the title and abstract of each paper, checking for inclusion and exclusion criteria. After this screening a second check was performed in which entire texts were scanned to ensure the papers were eligible for the review. The following information was extracted and entered into an Excel table for all included papers:

1. Author(s)
2. Title of Paper
3. Year
4. Academic Field
5. Definition of safety
6. Method

All but one (Savage. 1993) of the papers can be separated into using one of two types of methods: contingent valuation or discrete choice experiments (DCE). Therefore, two separate tables (Table 7 and Table 8) were then made for each type of method with columns for:

7. Paper

8. Scenario Description
9. Question asked to respondents
10. Measurement scale (CV) or Attributes (DCE)
11. Econometric Model(s)
12. Covariate results

The comprehensive search yielded a total of 679,467 results. Because the search terms ‘value’ and ‘review’ produced many seemingly irrelevant results, any results using these search terms were not included in the abstract screening, leaving 963 results for further screening. This first of all involved evaluating whether paper titles appeared to fit the inclusion criteria. This already excluded a large number of papers, namely 879 (91%). If a paper title was relevant then the abstract was checked to confirm that the paper did indeed fit the criteria. This was not always the case and thus from this first screening 49 papers (5%) were extracted. The bibliography of each of these papers was then scanned for reference titles that also referred to the empirical valuation of safety. 9 additional papers were found from the bibliography scans, meaning that 58 papers were eventually left for further screening. After the more thorough check, 24 of the 58 papers turned out to either be a non-empirical paper or to only mention the value of safety briefly, rather than it being the main crux of the paper. One additional paper was excluded as it only measured relative values of safety rather than absolute, using a ranking method (Savage. 1993). Therefore, in the end 33 papers were included and summarized in the review.

The main aim of this review, as mentioned previously, was to examine the various methodologies used for valuing safety. Therefore, in both the table and the findings section of this paper, the most weight will be placed on study methodology. Due to the variety of topics covered by the papers, the comparison of WTP values seemed nonsensical, however to give some insight into possible results from similar studies the covariate results that can be compared across fields are discussed in the findings.

4.1.3 Findings

Table 6 shows general information about the papers extracted from the review process. Regarding the fields of the papers, the most popular field is Environment (39%), followed by Transportation (21%) and Health (15%). Twenty-two of the papers (67%) used contingent valuation (CV) method for their valuation of safety and 11 (33%) used a form of discrete choice experiment (DCE) or conjoint analysis. Of the 33 papers, 20 (60%) used ‘risk reduction’ as the definition of safety, seven (21%) simply referred to a ‘reduction in [unwanted outcome]’, five papers (15%) used the term ‘safety’, and one paper (3%) valued ‘security’.

TABLE 6 GENERAL PAPER INFORMATION

Author(s)	Title of Study	Year	Academic Field	Definition of Safety	Elicitation Format
Alberini et al. (Alberini et al. 2006)	Willingness to Pay to Reduce Mortality Risks: Evidence from a Three-Country Contingent Valuation Study	2006	Health	Risk reduction	Contingent Valuation
Andersson (Andersson. 2007)	Willingness to pay for road safety and estimates of the risk of death: Evidence from a Swedish contingent valuation study	2012	Transport	Risk reduction	Contingent Valuation
Atkinson et al. (Atkinson et al. 2005)	Valuing the costs of violent crime: A stated preference approach	2015	Crime	Incidence reduction	Contingent Valuation
Carlsson et al. (Carlsson et al. 2004)	Is Transport Safety More Valuable in the Air?	2004	Transport	Risk reduction	Contingent Valuation
Carlsson & Johansson-Stenman (Carlsson and Johansson-Stenman. 2000)	Willingness to pay for improved air quality in Sweden	2000	Environment	Incidence reduction	Contingent Valuation
Carson & Mitchell (Carson and Mitchell. 1993)	The Value of Clean Water: The Public's Willingness to Pay for Boatable, Fishable, and Swimmable Quality Water	1993	Environment	Incidence reduction	Contingent Valuation

Author(s)	Title of Study	Year	Academic Field	Definition of Safety	Elicitation Format
Chanel et al. (Chanel et al. 2006)	Does public opinion influence willingness-to-pay? Evidence from the field	2006	Environment	Risk reduction	Contingent Valuation
Corso et al. (Corso et al. 2013)	A Comparison of Willingness to Pay to Prevent Child Maltreatment Deaths in Ecuador and the United States	2013	Health	Incidence reduction	Contingent Valuation
Dealy et al. (Dealy et al. 2013)	The Economic Impact of Project MARS (Motivating Adolescents to Reduce Sexual Risk)	2013	Health	Risk reduction	Contingent Valuation
Determann et al. (Determann et al. 2014)	Acceptance of Vaccinations in Pandemic Outbreaks: A Discrete Choice Experiment	2014	Health	Incidence reduction	Discrete Choice Experiment
Dickinson & Paskewitz (Dickinson and Paskewitz. 2012)	Willingness to Pay for Mosquito Control: How Important Is West Nile Virus Risk Compared to the Nuisance of Mosquitoes?	2012	Environment	Incidence reduction	Conjoint Analysis
Enneking(Enneking. 2004)	Willingness-to-pay for safety improvements in the German meat sector: the case of the Q&S label	2004	Food Safety	Safety	Discrete Choice Experiment

Author(s)	Title of Study	Year	Academic Field	Definition of Safety	Elicitation Format
Flügel et al. (Flügel et al. 2015)	Car drivers' valuation of landslide risk reductions	2015	Environment	Risk reduction	Discrete Choice Experiment
Garza-Gil et al. (Garza-Gil et al. 2016)	Marine aquaculture and environment quality as perceived by Spanish consumers. The case of shellfish demand	2016	Environment	Safety	Contingent Valuation
Georgiou et al. (Georgiou et al. 1998)	Determinants of individuals' willingness to pay for perceived reductions in environmental health risks: a case study of bathing water quality	1998	Environment	Risk reduction	Contingent Valuation
Gerking, et al. (Gerking et al. 1988)	The marginal value of job safety: A contingent valuation study	1998	Labour	Risk reduction	Contingent Valuation
Gyrd-Hanssen et al. (Gyrd - Hansen et al. 2008)	Willingness-to-pay for a statistical life in the times of a pandemic	2007	Health	Risk reduction	Contingent Valuation
Haddak et al. (Haddak et al. 2016)	Willingness-to-pay for road safety improvement	2014	Transport	Risk reduction	Contingent Valuation

Author(s)	Title of Study	Year	Academic Field	Definition of Safety	Elicitation Format
Halvorsen (Halvorsen. 1996)	Ordering effects in contingent valuation surveys: willingness to pay for reduces health damage from air pollution	1996	Environment	Risk reduction	Contingent Valuation
Henson (Henson. 1996)	Consumer Willingness to Pay for Reductions in the Risk of Food Poisoning in the UK	1996	Food Safety	Risk reduction	Contingent Valuation
Hunter et al. (Hunter et al. 2012)	The effect of risk perception on public preferences and willingness to pay for reductions in the health risks posed by toxic cyanobacterial blooms	2012	Environment	Risk reduction	Contingent Valuation
Iraguen & de Dios Orutzar (Iragüen and de Dios Ortúzar. 2004)	Willingness-to-pay for reducing fatal accident risk in urban areas: an Internet-based Web page stated preference survey	2004	Crime	Risk reduction	Discrete Choice Experiment
Khan et al. (Khan et al. 2014)	Household's willingness to pay for arsenic safe drinking water in Bangladesh	2014	Environment/Health	Risk reduction	Contingent Valuation
Loureiro & Umberger (Loureiro and Umberger. 2007)	A choice experiment model for beef: What US consumer responses tell us about relative preferences for food safety, country-of-origin labeling and traceability	2007	Food Safety	Safety	Discrete Choice Experiment

Author(s)	Title of Study	Year	Academic Field	Definition of Safety	Elicitation Format
Mattea et al. (Mattea et al. 2016)	Valuing landslide risk reduction programs in the Italian Alps: The effect of visual information on preference stability	2016	Environment	Risk reduction	Discrete Choice Experiment
Mofadal et al. (Mofadal et al. 2015)	Analysis of pedestrian accident costs in Sudan using the willingness-to-pay method	2015	Transport	Risk reduction	Contingent Valuation
Patil et al. (Patil et al. 2016)	Public preference for data privacy - A pan-European study on metro/train surveillance	2016	Transport	Security	Discrete Choice Experiment
Pham et al. (Pham et al. 2008)	Households' willingness to pay for a motorcycle helmet in Hanoi, Vietnam	2008	Transport	Incidence reduction	Contingent Valuation
Rizzi & Ortuzar (Rizzi and de Dios Ortúzar. 2003)	Stated preference in the valuation of interurban road safety	2003	Transport	Safety	Discrete Choice Experiment
Smith et al. (Smith et al. 2014)	How Should the Health Benefits of Food Safety Programs Be Measured?	2014	Food Safety	Risk reduction	Discrete Choice Experiment
Viscusi (Viscusi. 2009)	Valuing risks of death from terrorism and natural disasters	2009	Environment	Risk reduction	Discrete Choice Experiment

Author(s)	Title of Study	Year	Academic Field	Definition of Safety	Elicitation Format
Yabe (Yabe. 2016)	Students, Faculty, and Staff's Willingness to Pay for Emergency Texting	2016	Crime	Safety	Contingent Valuation
Yun et al. (Yun et al. 2016)	Analysis of the Relationship between Risk Perception and Willingness to Pay for Nuclear Power Plant Risk Reduction	2016	Environment	Risk reduction	Contingent Valuation

Table 7 synthesizes the more specific results of the papers that use CV methods. All papers use one of three types of measurement scale: open-ended questions, payment cards or dichotomous choice questions. Dichotomous choice questions can then be broken down into single- or double-bounded questions, where a double-bounded question means that after being given an initial 'yes or no' WTP price, as in a single-bounded question, the respondent is then given a second WTP option dependent on his first answer (Hanemann et al. 1991). The most popular question format of the 22 papers is open-ended questions (Carlsson and Johansson-Stenman. 2000, Chanel et al. 2006, Dealy et al. 2013, Georgiou et al. 1998, Gyrd - Hansen et al. 2008, Haddak et al. 2016, Halvorsen. 1996, Henson. 1996, Pham et al. 2008) (48%), followed by dichotomous choice (Alberini et al. 2006, Chanel et al. 2006, Corso et al. 2013, Garza-Gil et al. 2016, Khan et al. 2014, Pham et al. 2008, Yabe. 2016, Yun et al. 2016) (35%), and payment card (Atkinson et al. 2005, Carson and Mitchell. 1993, Gerking et al. 1988, Hunter et al. 2012, Mofadal et al. 2015) (22%). Two of the papers use both open-questions and dichotomous choice (Chanel et al. 2006, Pham et al. 2008). Of the six papers using dichotomous choice, two use double-bounded questions (Corso et al. 2013, Khan et al. 2014).

Table 7 also includes findings concerning covariates and their effect on WTP for safety. These covariates can be categorised into three groups: individual characteristics, individual relationship with risk, and aspects of the study design. First, we can consider variables that are strictly related to individual characteristics. Higher income is associated with a higher WTP in every case in which it is considered (Andersson. 2007, Atkinson et al. 2005, Carlsson et al. 2004, Georgiou et al. 1998, Gerking et al. 1988, Gyrd - Hansen et al. 2008, Haddak et al. 2016, Halvorsen. 1996, Henson. 1996, Hunter et al. 2012, Khan et al. 2014, Mofadal et al. 2015, Pham et al. 2008, Yun et al. 2016). In many of the papers having a higher level of education appears to have a positive effect on WTP (Alberini et al. 2006, Atkinson et al. 2005, Carlsson and Johansson-Stenman. 2000, Georgiou et al. 1998, Halvorsen. 1996, Mofadal et al. 2015, Pham et al. 2008), however there are also papers which show a higher level of education leading to a lower WTP (Gerking et al. 1988, Henson. 1996, Yun et al. 2016). Age and gender are variables that have ambiguous effects; several papers find that increasing age is associated with increased WTP (Alberini et al. 2006, Gerking et al. 1988, Gyrd - Hansen et al. 2008, Mofadal et al. 2015, Pham et al. 2008, Yabe. 2016), however some also produce the opposite result (Andersson. 2007, Carlsson and Johansson-Stenman. 2000, Halvorsen. 1996, Henson. 1996, Yun et al. 2016). In papers where gender is considered sometimes men report a higher WTP (Carlsson and Johansson-Stenman. 2000, Mofadal et al. 2015) and sometimes women do (Carlsson et al. 2004, Gyrd - Hansen et al. 2008, Henson. 1996).

TABLE 7 CONTINGENT VALUATION METHOD

Paper	Scenario Description	CV Question(s) asked to respondents	Measurement Scale	Econometric Model(s)	Covariate results
Crime					
Atkinson et al.	Respondents shown injury descriptions for 3 types of assault: common assault, other wounding, serious wounding. Also informed of pre-policy risk of the incident occurring.	Asked WTP to reduce chance of being a victim to one of the three offences (randomized per respondent) by 50% over the next 12 months. Payment vehicle is a one-off increase in local changes for law enforcement.	Payment card: £0-5000	Interval data model.	Severity of the risk increases WTP. Higher incomes and education both increase WTP.
Corso et al.	Respondents are asked to imagine that there is a program available in their city that reduces the risk of a child being killed by a parent or caretaker by 50%.	Asked WTP for this program through (i) taxes or (ii) donations.	Double-bounded dichotomous choice: Initial WTP value between $10 and $300. Second WTP values are $25 higher (lower) if response is 'yes' ('no').	Maximum likelihood function.	Those reporting history of child maltreatment have lower WTP.

Paper	Scenario Description	CV Question(s) asked to respondents	Measurement Scale	Econometric Model(s)	Covariate results
Yabe	Respondents are told that a Text-to-911 service would be paid for by a few charges to students, staff and faculty at the university.	Respondents are asked if they would be willing to pay X$ for an emergency text messaging service.	Dichotomous choice: bids - $1, $2, $3, $5, $10	Logit model.	Being interested in emergency texting, having experience in campus emergencies, being older, having a higher income, and being American (rather than international) leads to higher WTP.

Environment					
Carlsson & Johansson-Stenman	No scenario given, researchers want respondents to judge the information about air pollution from various sources.	Asked WTP for a 50% reduction in concentration of harmful substances where they live and work.	Open-ended questions.	Probit. Tobit type I. Tobit type II. Independent models.	WTP increases in income, wealth and education. WTP is larger for: men, members of environmental organizations, people living in big cities, and those who own their house or apartment. WTP is lower for retired people.

Carson & Mitchell	Respondents told that although present minimal water level is 'boatable', most of the nation's freshwater bodies are fishable and 70% are swimmable. Used the 'Resources for the Future' water quality index to show physical water quality parameters.	Asked WTP in taxes 'to keep the nation's freshwater bodies from falling below the boatable/fishable/swimmable level where they are now'. Four WTP amounts solicited for each of the three water quality questions: (i) amount given for each of the WTP questions (ii) WTP given after first amount is repeated and respondents encouraged to make desired corrections (iii) WTP after respondents informed of the range of the amounts households in their income group were already paying for water quality (iv) WTP given after respondents pushed to increase their bid.	Payment card: 0$ to a 'very high amount'. Five points on the card show average amounts households pay in taxes for non-environmental public goods.	Cobb-Douglas.	N/A
Chanel et al.	Respondents are given the hypothetical choice to move with their family to a less polluted city. Two cities are proposed that are equivalent with the exception of level of pollution and the cost of living.	Four steps: (i) WTP to live in less polluted city. (ii) WTP after shown mean WTP of all respondents. (iii) WTP after receiving scientific and quantitative information on health effects of pollution. (iv) WTP after new mean shown to respondents.	Dichotomous choice questions. Open-ended questions.	Wilcoxon sign-ranked tests.	Public opinion has no effect on WTP. Information provided leads to higher WTP.

Garza-Gil et al.	No scenario provided.	Asked WTP for an enhanced safety guarantee programme for shellfish quality and environmental conditions.	Dichotomous choice questions. WTP 5%, 10%, 20% or more than 20% more than normal price.	N/A	The higher the price the lower the number of people WTP for the intervention.
Georgiou et al.	Respondents informed about sewage contamination of bathing water and health risks from bathing, EC bathing water standards and actual quality of water at a particular beach.	Asked WTP for (i) for a gain (ii) for a loss in bathing water standards - dependent on the beach at which applicants are surveyed.	Open questions.	Semilog model.	Higher income and education lead to a higher WTP. The more unacceptable the respondent finds the risk, the higher their WTP. Having a family member who has suffered due to poor bathing water leads to a higher WTP.

Halvorsen	Respondents given description of benefits from a 50% reduction in air pollution. These are (i) reduction in in the risk of becoming ill and (ii) a reduction in damage due to acid rain.	Asked maximum WTP for 50% reduction in air-pollution. Four sub-samples with two splits: (1) Sub-samples B and D are told that the government will subsidize electric cars. A and C are told that the government uses a package of unspecified tools. (2) Those in A and B are given all information, then asked WTP. Those in C and D are first given health effect information then asked WTP, then are given all other effects and asked if they wish to change their WTP.	Open questions.	Tobit. Cragg (i) Probit model (ii) Truncated model.	Income, living in a major city, having a university degree and being concerned with the environment all have a positive effect on WTP. Age has a negative effect on WTP.
Hunter et al.	Respondents informed about (i) what cyanobacteria are (ii) the ecological and human health problems they cause and (iii) the practical options available for health-risk mitigation at Loch Leven.	Asked max WTP towards measures to reduce number of Risk Days from 90 to (i) 45 or (ii) 0. Payment vehicle is the cost of domestic water supply set by the council.	Payment card (values not stated).	Binary logit model. Non-parametric models: normal, logistic, lognormal, Weibull, and spike model.	Those with higher concern for environmental health risks have higher WTP. Income has a positive effect on WTP.

Khan et al.	No scenario provided.	Asked WTP for: (i) a communal deep tube well, (ii) one-time-off capital investment costs of the well (iii) one-time-off investment costs and (iv) operation and maintenance costs.	Double bounded dichotomous choice. Capital costs: Min. bid - 50 BDT. Max. Bid - 250 BDT. O&M costs: min. bid - 10 BDT. Max. Bid - 100 BDT.	Bivariate probit model. Random effects probit.	When respondents are male or earn higher incomes WTP is higher. If households are exposed to higher risk levels, if respondents are aware that their water is contaminated, and if household members are affected by arsenic exposure then WTP increases.
Yun et al.	Respondents are first asked to rank an image about nuclear power plants on a Likert scale of 'very good image/safe (5)' to 'very bad image/unsafe (1)'.	Respondents are asked if they would pay *A$* to reduce NPP hazard.	Dichotomous choice. Bids are not described.	Log-linear. Linear. Linear-log. Power regression models.	Higher scientific background/ low risk perception led to a lower mean WTP. Mean WTP decreased with increasing quality of informational image.
Food Safety					

Henson	Respondents are informed that chicken/egg consumption can cause food poisoning. They are told about two brands of chicken/eggs in the shop. Brand A and Brand B are identical except Brand A has been thoroughly tested and thus has a lower risk of giving one food poisoning.	Maximum additional amount WTP for a risk reduction in (i) fatal food poisoning (ii) mild food poisoning (iii) moderate food poisoning (iv) severe food poisoning in chicken or eggs.	Open question.	Ratios calculated.	More severe outcome leads to higher WTP. Personal experience of food poisoning has a negative effect on WTP. Mean WTP is higher for female respondents. Age and education both have a negative effect on WTP. WTP is positively affected by income.
Health					
Alberini et al.	Respondents are shown their baseline risk of death (that varies with gender and age) over the next 10 years.	Asked WTP for a risk-reduction in death of (i) 5-in-1000 incurred over the next 10 years, (ii) 1-in1000 incurred over the next 10 years, (iii) 5-in-1000 that begins at age 70 and is spread over next 10 years. Payment would be made every year.	Dichotomous choice questions.	Accelerated-life Weibull model.	Income increases WTP. WTP increase with age until age 60 and then plateaus. Hospitalization for cardiovascular or respiratory illness leads to higher WTP.

Dealy et al.	Participants are randomly assigned to receive one of three treatments: (i) sexual risk reduction intervention (ii) sexual risk reduction plus alcohol risk reduction (iii) sexual risk reduction intervention including both an alcohol and a marijuana risk reduction component.	Asked WTP 'not to get' (i) a curable STD (ii) an incurable non-fatal STD (iii) a fatal STD. Asked before and after intervention.	Open question with a bound of $0-100,000	Anova.	WTP increases after receiving the intervention. WTP increases with the severity of the STD.
Gyrd-Hanssen et al.	No scenario provided.	Asked maximum WTP in order to have a course of Tamiflu drug available in case they would need it.	Open questions.	Linear regression analysis.	Age and being female increase WTP. Household income has a positive impact on WTP. Being uncertain of baseline risk has a positive impact on WTP. Being uncertain of the perceived benefit has a negative impact on WTP.
Labour					

Gerking, et al.	Respondents are asked what their current job is.	Asked (i) how large an increase in annual wages would lead to respondent voluntarily working 'one step up' the risk ladder (WTA) (ii) how large a decrease in annual wages would a respondent forego in order to move one step lower on the risk ladder.	Payment card: $0 to $6000.	Two-limit tobit procedure.	Higher income and perceived likelihood of death at work leads to higher WTP/WTA. Older-workers have a higher WTP/WTA. WTP decreases with formal educational levels.
Transport					
Andersson	Respondents shown overall death risk for an individual. Also shown risk of dying in a traffic accident.	Asked one of two questions: (i) WTP for reducing personal annual risk of death by a third. (ii) WTP for reducing personal annual risk of dying in a traffic accident by one third.	Open-ended questions	Non-linear models. Log-linear models.	WTP increases as baseline risk increases. WTP declines with age. WTP declines with background risk. WTP increases with income.

Carlsson et al.	Two scenarios: (i) The respondent is going to take a taxi alone. They have two taxi options which are identical except for the risk of a fatal accident - 1 in 1 million (AAA) or 0.5 in 1 million (BBB). (ii) The respondent will take a plane alone. They have two airline options which are identical except for the risk of a fatal accident - 1 in 1 million (AAA) or 0.5 in 1 million (BBB).	Cases: Asked WTP for safer air trip compared to AAA at (i) 500 SEK (ii) 3000 SEK. Asked WTP for safer taxi ride compared to AAA at (iii) 50 SEK (iv) 500 SEK. (v) Asked both (i) and (iv). (vi) Asked both (ii) and (iv).	Open-ended questions anchored with baseline-risk prices (AAA)	Tobit type II.	Cost of trip leads to a higher WTP. Higher income leads to higher WTP. Male respondents have lower WTP. Fear of flying leads to a higher WTP for both air and taxi questions.
Haddak et al.	Three projects presented to respondents: reduces risk of being a victim of (i) a road accident that causes minor injuries (ii) a road accident resulting in serious injuries (iii) a road accident that results in moderate injuries.	Asked how much they would be willing to pay for a (i) 25% reduction (ii) 50% reduction in risk of experiencing various non-fatal types of injuries following a road accident.	Open questions.	Logit. Tobit.	WTP is higher for more severe injuries. WTP increases with income. Accidental experience of individuals (direct and indirect) leads to increased WTP.

Mofadal et al.	Respondent is told to imagine going to work or performing daily activities and during these they need to cross busy streets to reach their destination. The respondent can choose one of five options to reduce this risk.	The respondent first chooses the optimum scenario regarding: crossing behavior and side walking. They are then asked their maximum WTP to reduce the risk of a fatality in that scenario. They are also asked their maximum WTP for a pedestrian safety program that reduces fatality risk by 50%.	Payment card: 0 to more than 3000 SDP	Log-linear.	Age positively affects WTP. Income positively affects WTP. Married respondents have a lower WTP. Males have a higher WTP. Higher education increases WTP.
Pham et al.	Respondents are given the hypothetical situation that the government subsidizes the price of motorcycle helmets.	Respondents are asked the maximum amount they are willing to pay for a motorcycle helmet.	Open question. Dichotomous choice questions - min. 50,000 SDP, max. 150,000 SDP.	Interval regression. Multi-linear regression model.	Age and income have a positive effect on WTP. Those with higher education, those with jobs outside of the office and those with a better knowledge of/attitude towards helmets have a higher WTP.

Secondly, we can consider the group of variables that concern the individual and their relationship with the risk. For example, if an individual is more susceptible to the outcome (Alberini et al. 2006), has been previously exposed (Khan et al. 2014) to the outcome, or has a family member who has experienced the situation (Georgiou et al. 1998), they will report a higher WTP according to some of the papers reviewed. There are several other factors that could lead to an increased WTP: For example, if an individual is more concerned about the issue at risk (Halvorsen. 1996, Hunter et al. 2012), finds the risk unacceptable (Georgiou et al. 1998), has a higher perceived risk (Gerking et al. 1988, Gyrd - Hansen et al. 2008), is uncertain of the benefit or risk of the outcome (Gyrd - Hansen et al. 2008), or is aware of (Khan et al. 2014), interested in (Yabe. 2016), or knowledgeable about (Pham et al. 2008) the issue. Those with experience of the outcome sometimes report higher WTP (Dealy et al. 2013, Haddak et al. 2016, Yabe. 2016) and sometimes report lower WTP (Corso et al. 2013, Henson. 1996) than those who had not experienced the outcome. The papers where WTP is lower with experience of the outcome cover the topics of Child Maltreatment risk reduction (Corso et al. 2013) and the risk reduction of food poisoning (Henson. 1996). Corso et al. (2013) indicate that the finding is not what was expected, but they do not come up with a concrete explanation for the mechanism underlying the result. Henson (1996) explains his result through two mechanisms: the first is that those who have recently suffered from food poisoning believe that they have a smaller chance of getting food poisoning in the future, and the second is that many suffered only mild symptoms and so may underweight the probability of having moderate to severe food poisoning symptoms (Henson. 1996).

Thirdly, we can consider the group of variables related to aspects of the study design. Using a higher baseline risk (Andersson. 2007) or severity of risk (Atkinson et al. 2005, Dealy et al. 2013, Henson. 1996) is associated with individuals reporting a higher WTP. From the two CV studies that place a price on the intervention, one study finds that increased cost price is associated with higher WTP (Carlsson et al. 2004) while the other study finds the opposite result (Garza-Gil et al. 2016). Carlsson et al. (2004) give no explanation as to why having a more expensive taxi ride or flight is associated with a higher level of WTP, but it may be due to people assuming that the more expensive the journey is, the safer it is. Two studies also investigate the effects of more information on individuals' WTP. Chanel et al. (2006) find that giving people more information regarding pollution levels is associated with higher WTP, whereas Yun et al. (2016) find that providing people with better quality informational images is associated with lower WTP for reduced nuclear power plant hazard. Because they approach the study from the point of view that nuclear power plants are safer than assumed by some of the public, they do not explicitly discuss why it is that better quality information is associated with lower WTP (Yun et al. 2016).

As previously mentioned, the second most popular method for valuing safety is DCE or conjoint analysis. Table 8 summarizes the main traits of the papers in which DCE or conjoint analysis is used. The most obvious difference between DCE (or conjoint analysis) and CM methods is that DCE and conjoint analysis use attributes in order to extract the value being researched indirectly. Due to the papers in this review coming from many different fields, it is not possible to directly compare attributes. However, there are three types of attribute which practically every DCE paper uses and can be described in broad terms as: one which considers the cost price (Determann et al. 2014, Dickinson and Paskewitz. 2012, Enneking. 2004, Flügel et al. 2015, Loureiro and Umberger. 2007, Mattea et al. 2016, Patil et al. 2016, Rizzi and de Dios Ortúzar. 2003, Smith et al. 2014), one which considers the level of risk or risk reduction (Determann et al. 2014, Dickinson and Paskewitz. 2012, Flügel et al. 2015, Iragüen and de Dios Ortúzar. 2004, Loureiro and Umberger. 2007, Rizzi and de Dios Ortúzar. 2003, Smith et al. 2014, Viscusi. 2009), and one which considers the type of intervention (Determann et al. 2014, Dickinson and Paskewitz. 2012, Enneking. 2004, Flügel et al. 2015, Loureiro and Umberger. 2007, Mattea et al. 2016, Patil et al. 2016, Rizzi and de Dios Ortúzar. 2003, Smith et al. 2014, Viscusi. 2009).

TABLE 8 DISCRETE CHOICE EXPERIMENT / CONJOINT ANALYSIS

Paper	Scenario Description	Question asked to respondents	Attributes	Econometric Model(s)	Covariate Results
Crime					
Iraguen & de Dios Orutzar	Respondents are asked to imagine they are traveling to work from home. The trip takes place on a regular working day, they arrive at their destination at around 7:45 am, and they drive their own car and are responsible for all costs involved.	Respondents are asked to choose between two different routes with differing attributes.	Travel time. Travel cost. Number of fatal car accidents per year.	Multinomial logit model.	Income negatively affects the perception of the importance of travel cost. Safety valuation is positively affected if the individual travels with somebody else. This is also true if the respondent is female or they have been in a serious accident before.
Environment					

Paper	Scenario Description	Question asked to respondents	Attributes	Econometric Model(s)	Covariate Results
Dickinson & Paskewitz	Respondents are informed of the multiple types of mosquitoes in Madison (some a nuisance, some transmit West Nile virus). Control program which would control mosquito larvae could control one type of mosquito larvae or both.	Asked to choose between pairs of hypothetical control programmes.	West Nile Risk. Type of mosquito targeted. Cost (through taxes): $10-200.	Conditional logit model.	Increased risk level leads to higher WTP. WTP decreases as cost increases.
Flügel et al.	Respondents who had a recent trip by car were presented with different choices for a car trip route.	Respondents are asked to choose between two routes with four differing attributes, 6 times.	Cost: fuel and toll. Travel Time. Casualties: fatalities and serious injuries. Landslides: share of the route with landslide risk.	Mixed logit models.	Men are less likely to choose lower landslide risk. People with a higher education tend to choose the option with the lowest risk more often.

Paper	Scenario Description	Question asked to respondents	Attributes	Econometric Model(s)	Covariate Results
Mattea et al.	No scenario given.	Respondents are given six choice sets of seven alternatives, each of which consists of five attributes. These questions are asked twice, and respondents are given visual information on the possible options before being asked a second time.	Four alternatives represent devices to protect against landslides: Diverging channel, Retaining basin, Video cameras, and Acoustic sensors. The fifth alternative is a hypothetical road toll.	Mixed logit in WTP spaces. Multinomial logit model. Mixed logit in preference space.	The 'status quo' is negatively perceived.
Viscusi	Traffic - On an average day 100 people die due to traffic accidents. These risks are isolated deaths. Natural disasters - These are national catastrophes and large numbers of people can die at the same time. Hurricane Katrina killed over 1,000 people. Terrorism - Attacks by terrorists can also be catastrophes. The 9/11 attacks killed 2,976 people.	Respondents are asked 'risk-risk' tradeoff questions (traffic accident-terrorist attack, traffic accident-natural disaster) in two sets of 6 question blocks.	Type of deaths prevented. Average number of deaths prevented.	Conditional logit models. Mixed logit models	More education raises the utility coefficient in every instance, and more so with terrorism. Income has a negative effect on utility. Seatbelt usage increases the utility of reducing all deaths.

Paper	Scenario Description	Question asked to respondents	Attributes	Econometric Model(s)	Covariate Results
Food Safety					
Enneking	Participants are given a short introduction to the Quality and Safety labelling system (regarding liver sausages).	Asked to name three choices from a set of 6 sausages	Brand A: national premium brand (with/without Q&S label). Brand B: National brand (with/without label). Brand C: National premium brand - reduced fat (no label). Brand D: Private label - organic (no label). Brand E: National organic umbrella brand name (no label). Brand F: Private label.	Maximum likelihood.	Those who find low prices important avoid the more expensive labelled brands.
Loureiro & Umberger	No scenario given.	Asked to choose between two steaks (Option A and Option B) with five varying attributes.	Price ($/lb). Country of origin labeled. Traceability to the farm. Food safety inspected. Guaranteed tender.	Multinomial conditional logit model.	Increasing price of option leads to lower utility. Steaks inspected by US food inspectors carry the highest premium.

Paper	Scenario Description	Question asked to respondents	Attributes	Econometric Model(s)	Covariate Results
Smith et al.	No scenario given.	Regarding improvement of food safety respondents are asked to choose between: the 'status quo', 'Hire more inspectors' and 'purchase medicine'. Each subject is asked 12 choice questions, where each option consists of five attributes.	Annual risk of food borne illness. Average amount of time you will be sick. Extra time needed to prepare food. Cost. Annual increase in income tax.	Multinomial logit models.	Consumers prefer reduction ex ante risk than ex post. Those who are more willing to accept risk, are not as likely to accept risk reduction policies. Respondents prefer private control over the risk reduction.
Health					
Determann et al.	Respondents are presented with some combination of two scenario variables (i) susceptibility to the disease (ii) severity of the disease.	Asked to choose between Vaccine A, Vaccine B and 'No Vaccine' in 16 choice sets. Vaccines are comprised of different levels of 5 attributes.	Effectiveness of vaccine. Safety of the vaccine. Advice regarding the vaccine. Media coverage. Out-of-pocket costs.	Latent class model.	Females and individuals who stated they would never get vaccinated were more influenced by media and more sensitive to costs. WTP is higher for more effective vaccines, especially if the outbreak was more serious.
Transport					

Paper	Scenario Description	Question asked to respondents	Attributes	Econometric Model(s)	Covariate Results
Patil et al.	No scenario given.	Each respondent answered five choice exercises regarding their security preferences when traveling by train or metro.	Type of CCTV cameras. How long CCTV information is stored. Who can access CCTV information. Security personnel at the station. Type of security checks at the station. Time to go through security checks. Security surcharge.	Multinomial logit model.	All preferred CCTV over no CCTV. Preference is weaker for younger people. Females have a stronger preference for CCTV.
Rizzi & Ortuzar	Survey is disguised as a survey to improve interurban route policy and road safety. Respondents are given an identical trip in which: they drive their own car, they pay for the total cost of the trip, and they have to return after 20:00.	Respondents are asked to answer nine choice situations. They are asked to choose between two routes with differences in the three attributes.	Travel time. Toll charge. Annual accident rate (represents "general level of safety").	Binary logit models.	Women have a higher preference for safety than men, as do older people. There is a higher preference for safety if the trip takes place at night. A person driving with others in the car is more aware of risk.

Looking at the results from the DCE papers, the effects of covariates on WTP can, once again, be split into three groups – personal characteristics, individual relationship with risk, and aspects of the study design. From Table 8 we can see that higher age (Patil et al. 2016, Rizzi and de Dios Ortúzar. 2003), education (Flügel et al. 2015) and income (Iragüen and de Dios Ortúzar. 2004) all increase WTP. The only personal variable that differs from the CV results is that women (Determann et al. 2014, Flügel et al. 2015, Iragüen and de Dios Ortúzar. 2004, Rizzi and de Dios Ortúzar. 2003) always report a higher WTP. Regarding the interaction of individuals and risk; experience of the event (Iragüen and de Dios Ortúzar. 2004) is associated with higher WTP. Finally, looking at the variables which relate to the effectiveness of the method: a higher cost price is associated with lower WTP (Dickinson and Paskewitz. 2012, Loureiro and Umberger. 2007), while a more severe outcome (Atkinson et al. 2005), a higher risk level (Dickinson and Paskewitz. 2012) and a more effective treatment (Determann et al. 2014) are all associated with higher WTP.

Many of the papers in the study consider some theoretical issues that come with the methodology used. Out of the CV papers, most of those that do consider theory look at the use of visual aids to represent risk (Alberini et al. 2006, Andersson. 2007, Atkinson et al. 2005, Carlsson et al. 2004, Carson and Mitchell. 1993, Gerking et al. 1988). Other issues considered are sample size limitations (Pham et al. 2008, Yabe. 2016), embedding effects (Carlsson et al. 2004, Georgiou et al. 1998, Halvorsen. 1996), the interpretation of risk (Gyrd‐Hansen et al. 2008, Mofadal et al. 2015), and interviewing effects (Hunter et al. 2012). The most commonly considered theoretical issues in the DCE papers are sample bias (Dickinson and Paskewitz. 2012, Iragüen and de Dios Ortúzar. 2004), the use of visual aids (Mattea et al. 2016) and behaviour comparability (Loureiro and Umberger. 2007, Rizzi and de Dios Ortúzar. 2003). This being said, it is often in all papers that while they mention relevant theoretical problems, they do little to mitigate them in the research.

4.1.4 Discussion

This review aimed to synthesize the methodology and study design used in empirical research valuating safety. This is an issue which is becoming more and more relevant as CBA use increases for the evaluation of governmental policy, and as global threats become increasingly apparent across the globe. As can be seen from the results section above, there are several main findings regarding the valuation of safety. Firstly, the two main methods used are CV and DCE (or conjoint analysis), with CV being the most popular. Secondly, most papers use 'risk reduction' as a definition of safety when valuating it. Thirdly, we see that there are several covariate results that are measured across papers, all of which fell under three categories: individual characteristics, the relationship between the individual and risk, and aspects of the study design. Overall, it was the covariate results related to individual characteristics that led to the most ambiguous conclusions, while the results concerning the individual's relationship with risk mostly ran in the same direction across papers. Finally, while most papers did *mention* at least one of the theoretical issues discussed in the Theoretical Background section, few attempted to *tackle* the issues they mention.

Something that it not directly discussed in the findings but is noteworthy, is that all papers use an individual perspective when valuating safety, and none consider or mention using a societal perspective. Doing this would allow the measurement of how individuals value the safety of others and not just themselves, which is clearly relevant when policies are designed to improve the safety of citizens in general, and use taxes as the payment vehicle. Using a societal perspective in the methodological design would involve additional scenario description and questions. For example, one can include information in the scenario description about who is at risk and who benefits from the intervention, and also ask questions about the individual's WTP if others are also paying (e.g. through raising taxes), or if the individual themselves does or does not benefit (i.e., distinguishing between social values that do or do not take self-interest into account (Bobinac et al. 2013, Dolan et al. 2003)).

Several further observations can be made from this literature review. Firstly, there is the limited number of papers retrieved from the literature search. Therefore, it is difficult to make strong conclusions or recommendations from any of the results, especially those stemming from DCE experiments, of which there are relatively few. It would also be irresponsible to comment on similarities in methodologies used within fields, as the number of papers per field is too small to glean substantial evidence. Secondly, there is the complexity to defining safety. Even though most papers define safety as 'risk reduction' when valuating it, not all do, and so this muddles any comparison between papers that use different definitions. Lastly, there is the wide range of fields used in this research. Although the diversity of topics does show that the valuation of safety is relevant in many different areas, it is not useful for the comparison of results. These observations show us how useful a standardised method of safety valuation would be. Governments are presented many policy options while they have a restricted budget. Consequently, they must make choices about which policies to implement and which not, potentially concerning different departments, such as health and education. When making such choices, information about the value for money different policies generate is relevant information and in this context a standardised method for valuing safety would be beneficial for the comparability of information between policies.

The main limitation of this study is that it is not a systematic literature review. It would have increased the reliability of our results if a meta-analysis had been carried out. In a similar vein, the process could have been strengthened by a second author reviewing abstracts, the use of more than two databases, or the inclusion of more types of research (i.e. theses, papers in a language other than English). However, as this is the first literature review on safety valuation, the results still provide insight into an area of research that has not been heavily analysed.

Overall, it has become clear that there is little to no standardisation in safety valuation. Regarding which is 'the best' methodology to use, this literature review brings to light more questions than it does answers: What definition of safety is the best for its evaluation? Which stated preference method should be used, CV or DCE, and which methodological issues should be considered in study design? Should the individual or the societal view be applied in the context of valuing public goods? Which covariates should be added to gain the most insight into an individual's WTP? In other words, there still appears to be a long way ahead before consensus can be attained about a standardised methodology for valuating safety. In the meantime, forthcoming safety valuation research can build upon the findings of this review of the literature, and contribute to the development of more standardised methods by addressing questions about definition of safety, choice and design of method and perspective for valuation, and selection of covariates thoroughly and clearly.

5. Next steps

This section discusses the next steps to be carried out as part of Work Package 14.

5.1 Implementation of the case studies to measure and value elements of the system

Based on the work presented in this deliverable, the selection of the case studies and the case study methodology will be finalised, taking into account the results of a pilot case study. Data will be collected on the basis of a questionnaire provided to case study institutions, complemented by interviews with representatives of the institutions and field visits, where needed. In addition, complementary data will be collected from central data sources, where relevant. The data collected will be analysed and used to evaluate the cost-effectiveness of elements related to the system as described in Task 5 of Work Package 14.

5.2 Implementation of the methodology to measure and value elements related to the wider framework

Findings from the literature review presented in Section 4 will feed into the development of a questionnaire with the aim to estimate the value of safety relevant for the COMPARE project.

5.3 Next deliverables

The next deliverables to be produced by Work Package 14 are as follows:

- Deliverable 14.3. A scientific paper describing the methodology and results of estimating the value of safety, with the results from several European countries (to be submitted in Month 45 of the project);
- Deliverable 14.4. Report on the (potential) cost-effectiveness of COMPARE, based on the case studies (scenario/pilot/retrospective studies). Each case study presented will include a section on elements related to the system and on the wider framework (to be submitted in Month 54 of the project);
- Deliverable 14.5. A report on the assessment of options for refining selected elements of COMPARE in view of improving the overall cost-effectiveness of the system, with recommendations (to be submitted in Month 60 of the project).

References

Safety | Definition. http://www.merriam-webster.com/dictionary/safety?utm_campaign=sd&utm_medium=serp&utm_source=jsonld. Accessed 10/24 2016.

Alberini, A., Hunt, A., Markandya, A. (2006). Willingness to pay to reduce mortality risks: evidence from a three-country contingent valuation study. *Environmental and Resource Economics*, 33(2), 251-264.

Alpizar, F., Carlsson, F., Martinsson, P. (2003). Using choice experiments for non-market valuation. *Economic issues-stoke on trent*, 8(1), 83-110.

Andersson, H. (2007). Willingness to pay for road safety and estimates of the risk of death: Evidence from a Swedish contingent valuation study. *Accident Analysis & Prevention*, 39(4), 853-865.

Atkinson, G., Healey, A., Mourato, S. (2005). Valuing the costs of violent crime: A stated preference approach. *Oxford Economic Papers*, 57(4), 559-585.

Ball, D. J. (2000). Consumer affairs and the valuation of safety. *Accident Analysis & Prevention*, 32(3), 337-343.

Bleichrodt, H. & Eeckhoudt, L. (2006). Willingness to pay for reductions in health risks when probabilities are distorted. *Health Economics*, 15(2), 211-214.

Bobinac, A., Exel, N., Rutten, F. F., Brouwer, W. B. (2013). Valuing QALY gains by applying a societal perspective. *Health Economics*, 22(10), 1272-1281.

Boithias, L., Terrado, M., Corominas, L., Ziv, G., Kumar, V., Marqués, M., et al. (2016). Analysis of the uncertainty in the monetary valuation of ecosystem services—A case study at the river basin scale. *Science of the Total Environment*, 543, 683-690.

Brown, T. C. (2003). Introduction to stated preference methods. In *A primer on nonmarket valuation* (pp. 99-110): Springer.

Carlsson, F. & Johansson-Stenman, O. (2000). Willingness to pay for improved air quality in Sweden. *Applied Economics*, 32(6), 661-669.

Carlsson, F., Johansson-Stenman, O., Martinsson, P. (2004). Is transport safety more valuable in the air?. *Journal of Risk and Uncertainty*, 28(2), 147-163.

Carson, R. T. & Mitchell, R. C. (1993). The value of clean water: the public's willingness to pay for boatable, fishable, and swimmable quality water. *Water Resources Research*, 29, 2445-2454.

Cellini, S. R. & Kee, J. E. (2010). Cost-effectiveness and cost-benefit analysis. *Handbook of Practical Program Evaluation*, 3.

Chanel, O., Cleary, S., Luchini, S. (2006). Does public opinion influence willingness-to-Pay? Evidence from the field. *Applied Economics Letters*, 13(13), 821-824.

Chilton, S., Jones-Lee, M., Kiraly, F., Metcalf, H., Pang, W. (2006). Dread risks. *Journal of Risk and Uncertainty*, 33(3), 165-182.

Corso, P. S., Ingels, J. B., Roldos, M. I. (2013). A comparison of willingness to pay to prevent child maltreatment deaths in Ecuador and the United States. *International Journal of Environmental Research and Public Health*, 10(4), 1342-1355.

Dealy, B. C., Horn, B. P., Callahan, T. J., Bryan, A. D. (2013). The economic impact of project MARS (motivating adolescents to reduce sexual risk). *Health Psychology*, 32(9), 1003.

Determann, D., Korfage, I. J., Lambooij, M. S., Bliemer, M., Richardus, J. H., Steyerberg, E. W., et al. (2014). Acceptance of vaccinations in pandemic outbreaks: a discrete choice experiment. *PloS one*, 9(7), e102505.

Dickinson, K. & Paskewitz, S. (2012). Willingness to pay for mosquito control: How important is West Nile virus risk compared to the nuisance of mosquitoes?. *Vector-borne and Zoonotic Diseases*, 12(10), 886-892.

Dolan, P., Olsen, J. A., Menzel, P., Richardson, J. (2003). An inquiry into the different perspectives that can be used when eliciting preferences in health. *Health Economics*, 12(7), 545-551.

Enneking, U. (2004). Willingness-to-pay for safety improvements in the German meat sector: the case of the Q&S label. *European Review of Agricultural Economics*, 31(2), 205-223.

Flügel, S., Rizzi, L. I., Veisten, K., Elvik, R., de Dios Ortúzar, J. (2015). Car drivers' valuation of landslide risk reductions. *Safety Science*, 77, 1-9.

Garza-Gil, M. D., Vázquez-Rodríguez, M. X., Varela-Lafuente, M. M. (2016). Marine aquaculture and environment quality as perceived by Spanish consumers. The case of shellfish demand. *Marine Policy*, 74, 1-5.

Georgiou, S., Langford, I. H., Bateman, I. J., Turner, R. K. (1998). Determinants of individuals' willingness to pay for perceived reductions in environmental health risks: a case study of bathing water quality. *Environment and Planning A*, 30(4), 577-594.

Gerking, S., De Haan, M., Schulze, W. (1988). The marginal value of job safety: a contingent valuation study. *Journal of Risk and Uncertainty*, 1(2), 185-199.

Gyrd-Hansen, D., Halvorsen, P. A., Kristiansen, I. S. (2008). Willingness-to-pay for a statistical life in the times of a pandemic. *Health Economics*, 17(1), 55-66.

Haddak, M. M., Lefèvre, M., Havet, N. (2016). Willingness-to-pay for road safety improvement. *Transportation Research Part A: Policy and Practice*, 87, 1-10.

Halvorsen, B. (1996). Ordering effects in contingent valuation surveys. *Environmental and Resource Economics*, 8(4), 485-499.

Hanemann, M., Loomis, J., Kanninen, B. (1991). Statistical Efficiency of Double-Bounded Dichotomous Choice Contingent Valuation. *American Journal of Agricultural Economics*, 73(4), 1255-1263.

Henson, S. (1996). Consumer willingness to pay for reductions in the risk of food poisoning in the UK. *Journal of Agricultural Economics*, 47(1-4), 403-420.

Herzog Jr, H. W. & Schlottmann, A. M. (1990). Valuing risk in the workplace: market price, willingness to pay, and the optimal provision of safety. *The Review of Economics and Statistics*, 463-470.

Hunter, P. D., Hanley, N., Czajkowski, M., Mearns, K., Tyler, A. N., Carvalho, L., et al. (2012). The effect of risk perception on public preferences and willingness to pay for reductions in the health risks posed by toxic cyanobacterial blooms. *Science of the Total Environment*, 426, 32-44.

Iragüen, P. & de Dios Ortúzar, J. (2004). Willingness-to-pay for reducing fatal accident risk in urban areas: an Internet-based Web page stated preference survey. *Accident Analysis & Prevention*, 36(4), 513-524.

Kahneman, D. & Knetsch, J. L. (1992). Valuing public goods: the purchase of moral satisfaction. *Journal of Environmental Economics and Management*, 22(1), 57-70.

Khan, N. I., Brouwer, R., Yang, H. (2014). Household's willingness to pay for arsenic safe drinking water in Bangladesh. *Journal of Environmental Management*, 143, 151-161.

Logar, I. & van den Bergh, Jeroen CJM. (2012). Respondent uncertainty in contingent valuation of preventing beach erosion: An analysis with a polychotomous choice question. *Journal of Environmental Management*, 113, 184-193.

Loureiro, M. L. & Umberger, W. J. (2007). A choice experiment model for beef: What US consumer responses tell us about relative preferences for food safety, country-of-origin labeling and traceability. *Food Policy*, 32(4), 496-514.

Mattea, S., Franceschinis, C., Scarpa, R., Thiene, M. (2016). Valuing landslide risk reduction programs in the Italian Alps: The effect of visual information on preference stability. *Land Use Policy*, 59, 176-184.

Mitchell, R. C. & Carson, R. T. (1989). Using surveys to value public goods: the contingent valuation method: Resources for the Future.

Mofadal, A. I., Kanitpong, K., Jiwattanakulpaisarn, P. (2015). Analysis of pedestrian accident costs in sudan using the willingness-to-pay method. *Accident Analysis & Prevention*, 78, 201-211.

Patil, S., Patruni, B., Potoglou, D., Robinson, N. (2016). Public preference for data privacy–A pan-European study on metro/train surveillance. *Transportation Research Part A: Policy and Practice*, 92, 145-161.

Pham, K. H., Le Thi, Q. X., Petrie, D. J., Adams, J., Doran, C. M. (2008). Households' willingness to pay for a motorcycle helmet in Hanoi, Vietnam. *Applied Health Economics and Health Policy*, 6(2-3), 137-144.

Rizzi, L. I. & de Dios Ortúzar, J. (2003). Stated preference in the valuation of interurban road safety. *Accident Analysis & Prevention*, 35(1), 9-22.

Ryan, M., Gerard, K., Amaya-Amaya, M. (2007). *Using discrete choice experiments to value health and health care*: Springer Science & Business Media.

Savage, L. (1993). An empirical investigation into the effect of psychological perceptions on the willingness-to-pay to reduce risk. *Journal of Risk and Uncertainty*, 6(1), 75-90.

Sjöberg, L. (1998). Worry and risk perception. *Risk analysis*, 18(1), 85-93.

Smith, V. K., Mansfield, C., Strong, A. (2014). How should the health benefits of food safety programs be measured. *Advances in Health Economics and Health Services Research*, 24, 161-202.

Viscusi, W. K. (2009). Valuing risks of death from terrorism and natural disasters. *Journal of Risk and Uncertainty*, 38(3), 191-213.

Yabe, M. (2016). Students, Faculty, and Staff's Willingness to Pay for Emergency Texting. *Journal of Applied Security Research*, 11(4), 437-449.

Yun, M., Lee, S. H., Kang, H. G. (2016). Analysis of the Relationship between Risk Perception and Willingness to Pay for Nuclear Power Plant Risk Reduction. *Science and Technology of Nuclear Installations*, 2016.

Zeckhauser, R. (1996). The economics of catastrophes. *Journal of Risk and Uncertainty*, 12(2-3), 113-140.