DON'T READ THIS DISSERTATION

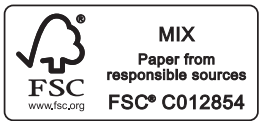
(Alternatively: choose to do so*)

Discrete Choice Experiments for Health Policy: Past, present, and future

By Domino Determann

Discrete Choice Experiments for Health Policy: Past, present, and future

Domino Determann



ISBN: 978-94-6233-349-9

Author: Domino Determann
Cover design: Bert Determann
Cover lay-out: Ontwerpbureau STUDIO|0404 - Nijmegen
Lay-out and printing: Gildeprint - Enschede

This PhD trajectory was funded by the European Union Seventh Framework Programme ((FP7/2007-2013) under grant agreement no. 278763), the Department of Public Health of Erasmus MC University Medical Center Rotterdam and by the National Institute for Public Health and the Environment (RIVM).

Printing of this dissertation was financially supported by the Department of Public Health of Erasmus MC University Medical Center Rotterdam and by the National Institute for Public Health and the Environment (RIVM).

Discrete Choice Experiments for Health Policy: Past, present, and future

Discrete keuze experimenten voor gezondheidsbeleid:
verleden, heden en toekomst

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de rector magnificus

Prof.dr. H.A.P. Pols

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
woensdag 28 september 2016 om 11.30 uur

door

Domino Determann
geboren te Rotterdam

Erasmus University Rotterdam

PROMOTIECOMMISSIE

Promotor:	Prof.dr. E.W. Steyerberg
Overige leden:	Prof.dr. J.H. Richardus Prof.dr. B. Donkers Prof.dr. A.M. Stiggelbout
Copromotoren:	Dr. E.W. de Bekker-Grob Dr. M.S. Lambooij

CONTENT

1

General introduction

Patient-centeredness, patient empowerment, shared decision-making and self-management are fundamental elements of current health policy in the Netherlands [1-3]. The changing role of the healthcare consumer to a more active one, is part of a broader movement away from the classical welfare state towards a participation society in which citizens are required to take responsibility for their own life, including health [4]. One concrete example is that citizens are, since the introduction of the universal health insurance scheme with managed competition in 2006 [2, 5, 6], expected to critically assess the available health plans annually, and to switch to another insurer's health plan if that better meets their preferences.

The changing role of the healthcare consumer does not only apply at the individual level (i.e. patient-physician and consumer-insurer relationship), but also at the collective level [1, 7]: the involvement of consumers in health policy decision-making is also being encouraged more and more. For example, the World Health Organization published a global strategy that advocates a shift towards people-centred[1] and integrated health services, in which more engagement is one of the core strategic directions [8]. Engagement is hereto defined as *"People and communities being involved in the design, planning and delivery of health services, enabling them to make choices about care and treatment options or to participate in strategic decision-making on how, where and on what health resources should be spent"*. Direct involvement of consumers in decision-making can be achieved by, for example, the participation of public or patient representatives in a decision committee during the health technology assessment process [9, 10]. Indirect involvement includes the elicitation of public perspectives by using, for example, focus group discussions and questionnaires [11]. By involving the public, health policies will better reflect public preferences [12]. Preferences can have an impact on the acceptance of and satisfaction with products, services or interventions, as well as on outcomes [13].

Although numerous research techniques are available to elicit public preferences for health-related topics, this dissertation focuses specifically on the discrete choice experiment (DCE) methodology to quantify public preferences. The main advantage of this methodology over other techniques, such as standard gamble or time trade-off, is that the DCE task more closely resembles the nature of real-life decision-making between products, services, or interventions [14, 15].

In this introductory chapter, relevant concepts are introduced first; the objectives as well as the content of the following chapters of this dissertation will be presented subsequently.

[1] People-centred care is not the same as patient-centred care [12]. The former reflects involvement of the public in all levels of health systems, including involvement in planning and developing of health services, while the latter relates to collaboration with the healthcare professional in individual clinical decision-making to ensure it reflects his/her own values and needs [50].

THE DISCRETE CHOICE EXPERIMENT METHODOLOGY

How many choices have you made today? One choice that is obvious at this point: you made the choice to read this dissertation. All choices involve at least two alternatives that need to be evaluated before making a choice: an alternative to reading this dissertation is *not* to read this dissertation, or to read another dissertation. There might be a number of reasons, or underlying preferences, that explain your choice. For example, the cover that attracted your attention, the topic is of interest, or you are curious to see what I wrote as you know me, and one reason might be more important for your choice than another reason. The goal of choice modelling is to better understand how choices are made, by quantifying the strength of underlying preferences, and to forecast future choice responses [16].

A DCE is a survey-based stated-preference elicitation technique: respondents are asked to state their preference for a product, service, or intervention, in hypothetical choice situations. Stated-preferences represent a hypothetical choice, while revealed preferences can be observed through choice behaviour in real market situations [16]. The main advantage of studying stated-preferences over revealed preferences, is that preferences can be elicited in a controlled situation. The DCE technique is particularly useful for products, services or interventions for which no market exists (yet). A DCE survey typically consists of a series of hypothetical choice situations (called choice sets), each concerning a discrete choice between two or more alternatives. The word 'discrete' indicates that respondents can only choose one of the alternatives. The presented alternatives are decomposed into characteristics (called attributes) to describe the alternative. Alternatives are distinguished from one-another by the systematic variation in the values of the characteristics (called attribute levels). See Figure 1 for an example of a choice set.

DCEs originate from mathematical psychology [18], and are, since pioneering work of Louviere and colleagues [19], standard practice in the field of marketing, transportation and environmental economics. The first DCE in health economics was published in 1990 by Propper [20]. Ryan and colleagues further introduced the method in health economics later in the 1990's [21-23]. Subsequently, the body of scientific literature is increasing rapidly (Figure 2).

Figure 1 | Example of a choice set concerning migraine treatment (adapted from [17]).

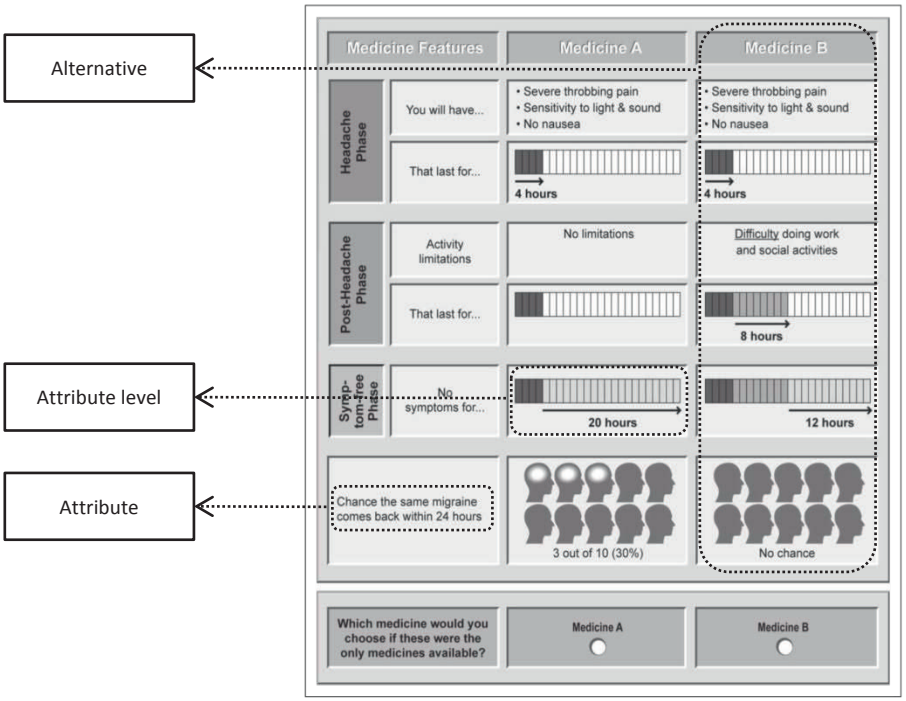
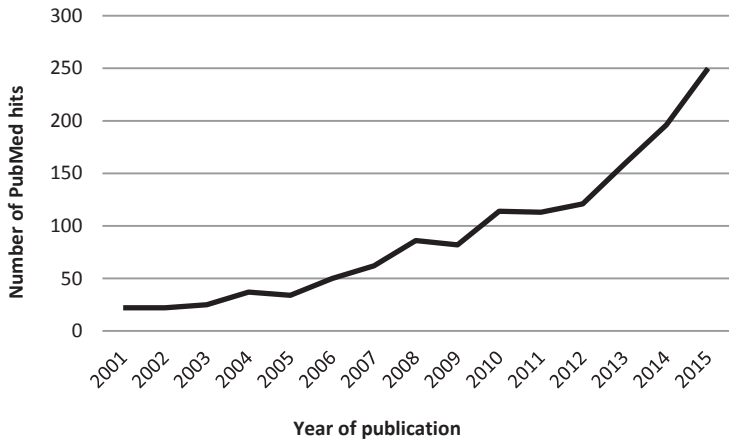


Figure 2 | Number of PubMed hits since 2001 ['discrete choice experiment' and equivalent terms], by year of publication.



THEORETICAL FOUNDATION OF DISCRETE CHOICE EXPERIMENTS

DCEs have their foundation in random utility theory (RUT)[2], first described by Thurstone in 1927 [24], and extended on by Manski [25] and Nobel Prize winner McFadden [26, 27]. They are consistent with Lancaster's economic theory of value, in which it is assumed that products, services, or interventions can be valued in terms of their constituent characteristics [28].

The true, but unobservable and therefore indirect, utility function *U* for alternative *i* of respondent *n* (U_{in}) consists of a systematic component that reflects the factors that are observable by the researcher (V_{in}), and a random component that reflects the factors that are not or cannot be observed by the researcher (ε_{in}) [16]:

[Equation 1]: $U_{in} = V_{in} + \varepsilon_{in}$

RUT assumes that respondent *n* evaluates all available information, and will choose alternative *i* within a choice set that yields the highest utility (or lowest regret in the case of random regret modelling [29]). The actual distribution of the random component of utility (ε_{in}) in the population is unknown; estimation of utility is therefore based on probabilistic choice modelling. The probability that respondent *n* prefers alternative *a* over alternative *b* in a particular choice set is expressed as follows [30]:

[Equation 2]:
$$\text{Prob}\,(U_{an} > U_{bn}\ \forall\ b{\neq}a) =$$
$$\text{Prob}\,(V_{an} + \varepsilon_{an} > V_{bn} + \varepsilon_{bn}\ \forall\ b{\neq}a) =$$
$$\text{Prob}\,(V_{an} - V_{bn} > \varepsilon_{bn} - \varepsilon_{an}\ \forall\ b{\neq}a)$$

The impact of the attributes on the systematic component of utility (V_{in}) is estimated based on the choices that respondents make in the DCE survey:

[Equation 3]: $V_{in} = \beta X_{in}$

Where X_{in} is the vector of the attributes for alternative *i* for respondent *n* and β[3] is the vector of coefficients to be estimated. The β for a certain attribute represents the weight of that attribute to overall utility for that alternative, which is sometimes called part-worth utility. The larger the β for one attribute, compared to another attribute, the more important this attribute is for someone's choice. The sign of the β reflects whether the effect on someone's choice is positive or negative.

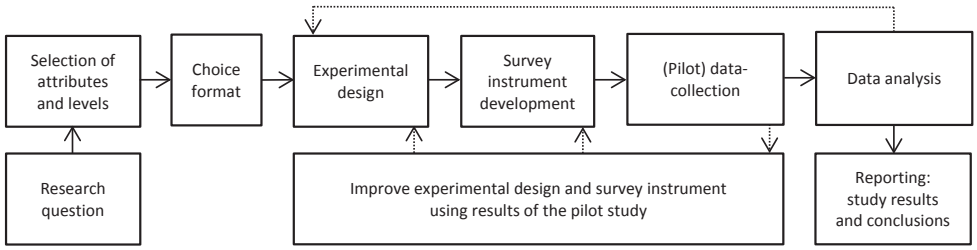
[2] Utility is the level of happiness that an alternative yields to an individual [16].
[3] The vectors of coefficients β are scaled with the variance of the unobserved factors (*λ)*, and are therefore only comparable in absolute terms within one model, not between models.

Based on these part-worth utilities, researchers can determine the relative importance of attributes, the trade-offs between attributes by calculating marginal rates of substitution (i.e. the degree to which respondents are willing to trade one attribute to the other), calculate total utility scores, and predict uptake rates.

HOW TO UNDERTAKE A DISCRETE CHOICE EXPERIMENT?

A number of guidelines have been written to support researchers when carrying out health-related DCEs [31, 32], including guidelines that focus on specific aspects of DCEs [33-35], and a guideline that focuses on conducting DCEs in low income countries [36]. Figure 3 presents a schematic overview of the scientific process of conducting a DCE. Each step is summarized briefly below.

Figure 3 | Schematic overview of the scientific process of conducting a DCE.



The first step in a DCE is the selection of attributes and attributes levels. Attributes are usually identified through literature study and qualitative research, for example expert interviews and/or focus group discussions [37]. Attributes generally need to be relevant for respondents as well as for the policy environment. DCEs can include only a limited number of attributes. Next, appropriate levels need to be assigned to the selected attributes. To avoid ignorance because of too small differences in levels, specification of a sufficient wide range of levels is necessary. Levels can include possible future levels that are currently unavailable. A minimum number of three attribute levels needs to be specified if non-linear effects are expected.

The following step is to define the choice format. The researcher needs to decide on whether to use labelled or unlabelled alternatives [38], and on the number of alternatives that will be presented within a choice set [39, 40]. In addition, the researcher needs to decide on either a forced or an unforced choice format. An unforced choice format involves the inclusion of an opt-out option (i.e. choosing not to buy the product, or not to use the service), a neither option (i.e. choosing neither of the alternatives), and/or a status quo option (i.e. choosing your current alternative).

Based on the utility function one wants to estimate, an experimental design needs to be constructed next. The number of all possible combinations of attributes and levels is often too large to be used in practice; a subset of all possible choice sets is therefore generated, usually using statistical software (e.g. Ngene or SAS). The design should enable the researcher to estimate unbiased estimates for every parameter in the model.

Once the design is created, the survey instrument can be developed. The presentation of the attributes and levels within a choice set, such as attribute ordering [41], attribute framing [42], and the use of either words or graphics [43] might have unintended influences on the obtained estimates, and therefore needs to be thoroughly considered. The survey instrument could further include an introductory text, practice choice sets, and general questions on respondent characteristics, health attitudes and health beliefs. In addition, it needs to be decided how the DCE survey will be administered: via pen-and-paper (either completed on-site, or distributed via mail) or electronically (either on-site using a laptop, or distributed via Internet).

The final data-collection is usually preceded by a pilot study. Pilot testing is necessary to detect instrument errors, as well as to test for respondent understanding. In addition, the parameter estimates of a pilot can serve as priors to improve the experimental design. The required sample size to being able to answer the research question also needs to be determined [35, 44].

Once the data is collected, econometric modelling techniques are used to estimate the part-worth utilities as described above. Several aspects determine which model is appropriate to analyse the DCE data: 1) the distributional assumption of the random component of utility (ε_{in}), 2) the choice set format (binary or multiple alternatives), 3) whether or not panel data was used (one respondent answering a number of choice sets), and 4) whether or not the analyst is interested in preference heterogeneity.

OBJECTIVES OF THIS DISSERTATION

This dissertation aims to contribute to the growing field of health-related stated-preference research by addressing research questions that relate to the past, the present and the future of the DCE methodology.

Part I: The past - review of the literature

1. What are recent practices and trends, including progress in methodology, in applications of health-related DCEs?

Part II: The present - three state of the art applications

2a. Stay healthy: What are preferences of European citizens for vaccination programmes during future pandemics?

2b. Anticipate future health: What trade-offs do consumers make between basic health plan characteristics in the Dutch health insurance market?

2c. Manage health and sickness: What are personal health record preferences of potential users?

Part III: The future - methodological studies

3a. Does the inclusion of an "opt-out" instead of a "neither" alternative affect DCE results? Does the inclusion of a "status quo" alternative in addition to an "opt-out" alternative affect DCE results in markets where there is a status quo?

3b. Does the choice of DCE administration mode (paper versus online) affect the result of a DCE?

3c. How do respondents complete the choice sets in a DCE?

OUTLINE OF THIS DISSERTATION

This dissertation consists of nine studies that relate to these research questions.

Part I of this dissertation (**Chapter 2**, research question 1) gives an overview of recent DCE practice, including progress in methodology, by systematically reviewing health-related DCEs published between 2009 and 2012. Having an overview of past DCE practices is important since this gives insights in methodological developments, and the current state of practice. In addition, such an overview provides directions for future research. Although systematic reviews of DCEs on specific health-related topics, for example health workforce policy [45] and pharmacy services [46], have been published recently, the latest general systematic review of health-related DCEs included studies published only up until 2008 [47-49]. A more recent general overview is lacking.

Part II of this dissertation presents a focus group discussion study and four DCE studies. The first three chapters of this part (**Chapters 3-5**, research question 2a) describe studies that were carried out in the context of the FP 7 funded project 'Effective Communication in Outbreak Management: development of an evidence-based tool for Europe (ECOM)'. **Chapter 3** describes a qualitative exploration of public opinions and attitudes for pandemics and vaccinations across three European countries. **Chapter 4** presents the results of the subsequently developed DCE on public preferences for vaccination programmes during pandemics in four European countries. **Chapter 5** further explores within-country differences in preferences for pandemic vaccinations by studying the DCE data of one of the included

European countries (the Netherlands) in more detail. **Chapter 6** concerns the second DCE application (research question 2b). It quantifies consumer trade-offs between health plan characteristics, and potential differences herein, according to age, health status and income, in the Dutch health insurance market with managed competition. The third DCE application is presented in **Chapter 7**, and focuses on potential user preferences for personal health records in the Netherlands (research question 2c).

Part III presents three methodological DCE studies. **Chapter 8** (research question 3a) focuses on the use of opt-out, neither and status quo alternatives in DCEs. The current guidelines do not request to assess the appropriate unforced choice format critically. If researchers do not select the unforced choice that best mimics the real market situation, welfare estimates could be biased. Evidence is needed to guide researchers in selecting the right unforced choice format in future DCEs. **Chapter 9** (research question 3b) describes if and how the mode of DCE administration (either via pen-and-paper or online) affects the outcomes of the DCE. Studies show that the presentation of attributes and levels might influence the results, it can be questioned whether the mode of administration likewise affect DCE results. This is especially important, since it is increasingly common to collect DCE data electronically, and more particularly, online. **Chapter 10** (research question 3c) describes an interview study that explores how respondents complete the choice sets of a DCE. In order to make valid inferences about individuals' preferences, it is essential that researchers understand how participants complete the choice sets in a DCE.

This dissertation concludes with the general discussion (**Chapter 11**), in which the main findings of the previous chapters are integrated and further discussed. Directions for future research are formulated. The chapter concludes with recommendations for health policy and recommendations for DCE researchers.

REFERENCES

1. [Good patientship]. De patiënt als sturende kracht. Den Haag: Council for Public Health and Health Care (RVZ), 2010. Available from: http://www.raadrvs.nl/uploads/docs/Advies_-_de_patient_als_sturende_kracht.pdf. *In Dutch.*
2. Schäfer W, Kroneman M, Boerma W, et al. The Netherlands: Health system review. 2010. Available from: http://www.euro.who.int/__data/assets/pdf_file/0008/85391/E93667.pdf?ua=1.
3. Ouwens M, Faber M, Van der Burg S, et al. [Shared Decision-making & Selfmanagement]. Shared Decision Making & Zelfmanagement. Nijmegen: Scientific Institute for Quality of Healthcare (IQ healthcare), 2012. Available from: https://www.raadrvs.nl/uploads/docs/Achtergrondstudie_Shared_DecisionMaking_en_Zelfmanagement.pdf. *In Dutch.*
4. [King's speech 2013]. Troonrede 2013. Den Haag: 2013. Available from: https://www.rijksoverheid.nl/documenten/toespraken/2013/09/17/troonrede-2013. [Last accessed 19 01 2016]. *In Dutch.*
5. Enthoven AC, van de Ven WP. Going Dutch--managed-competition health insurance in the Netherlands. NEJM. 2007;357:2421-3.
6. Health Insurance in the Netherlands. Den Haag: Ministry of Health, Welfare and Sports, 2011. Available from: http://www.government.nl/files/documents-and-publications/leaflets/2012/09/26/health-insurance-in-the-netherlands/health-insurance-in-the-netherlands.pdf.
7. Litva A, Coast J, Donovan J, et al. 'The public is too subjective': public involvement at different levels of health-care decision making. Soc Sci Med. 2002;54:1825-37.
8. WHO global strategy on people-centred and integrated health services - Interim report. Switzerland: World Health Organization (WHO), 2015. Available from: http://apps.who.int/iris/bitstream/10665/155002/1/WHO_HIS_SDS_2015.6_eng.pdf?ua=1.
9. Gagnon MP, Desmartis M, Lepage-Savary D, et al. Introducing patients' and the public's perspectives to health technology assessment: A systematic review of international experiences. Int J Technol Assess Health Care. 2011;27:31-42.
10. Bridges JF, Jones C. Patient-based health technology assessment: a vision of the future. Int J Technol Assess Health Care. 2007;23:30-5.
11. Rowe G, Frewer LJ. Public participation methods: A framework for evaluation. Sci Technol Hum Values. 2000;25:3-29.
12. Florin D, Dixon J. Public involvement in health care. BMJ. 2004;328:159-61.
13. Dirksen CD. The use of research evidence on patient preferences in health care decision-making: issues, controversies and moving forward. Expert Rev Pharmacoecon Outcomes Res. 2014;14:785-94.
14. Ryan M, Scott DA, Reeves C, et al. Eliciting public preferences for healthcare: a systematic review of techniques. Health Technol Assess. 2001;5:1-186.
15. Carson RT, Louviere JJ. A common nomenclature for stated preference elicitation approaches. Environ Resour Econ. 2011;49:539-59.
16. Hensher DA, Rose JM, Greene WH. Applied choice analysis: a primer. New York: Cambridge University Press; 2005.
17. Gonzalez JM, Reed Johnson F, Runken MC, et al. Evaluating migraineurs' preferences for migraine treatment outcomes using a choice experiment. Headache. 2013;53:1635-50.
18. Luce D, Tukey J. Simultaneous conjoint measurement: a new type of fundamental measurement. J Math Psychol. 1964;1:1-27.
19. Louviere JJ, Woodworth G. Design and analysis of simulated consumer choice or allocation experiments: an approach based on aggregate data. J Marketing Res. 1983:350-67.
20. Propper C. Contingent valuation of time spent on NHS waiting lists. Econ J. 1990:193-9.
21. Ryan M, Hughes J. Using conjoint analysis to assess women's preferences for miscarriage management. Health Econ. 1997;6:261-73.
22. Hanley N, Ryan M, Wright R. Estimating the monetary value of health care: lessons from environmental economics. Health Econ. 2003;12:3-16.

23. Ryan M, Farrar S. Using conjoint analysis to elicit preferences for health care. BMJ. 2000;320:1530-3.
24. Thurstone LL. A law of comparative judgment. Psychol Rev. 1927;34:273.
25. Manski CF. Structure of Random Utility Models. Theor Decis. 1977;8:229-54.
26. McFadden D. Conditional logit analysis of qualitative choice behavior. In: Zarembka P, editor. Frontiers of Economics. New York: Academic Press; 1974. p. 105-42.
27. The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2000. Available from: http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/2000/. [Last accessed 11 05 2016].
28. Lancaster KJ. A new approach to consumer theory. J Polit Econ. 1966;74:132-57.
29. De Bekker-Grob EW, Chorus CG. Random regret-based discrete-choice modelling: an application to healthcare. Pharmacoeconomics. 2013;31:623-34.
30. Train KE. Discrete choice methods with simulation. New York: Cambridge University Press; 2009.
31. Lancsar E, Louviere J. Conducting discrete choice experiments to inform healthcare decision making: a user's guide. Pharmacoeconomics. 2008;26:661-77.
32. Bridges JF, Hauber AB, Marshall D, et al. Conjoint analysis applications in health--a checklist: a report of the ISPOR Good Research Practices for Conjoint Analysis Task Force. Value Health. 2011;14:403-13.
33. Coast J, Al-Janabi H, Sutton EJ, et al. Using qualitative methods for attribute development for discrete choice experiments: issues and recommendations. Health Econ. 2012;21:730-41.
34. Reed Johnson F, Lancsar E, Marshall D, et al. Constructing experimental designs for discrete-choice experiments: report of the ISPOR Conjoint Analysis Experimental Design Good Research Practices Task Force. Value Health. 2013;16:3-13.
35. De Bekker-Grob EW, Donkers B, Jonker MF, et al. Sample Size Requirements for Discrete-Choice Experiments in Healthcare: a Practical Guide. Patient. 2015;8:373-84.
36. Mangham LJ, Hanson K, McPake B. How to do (or not to do) ... Designing a discrete choice experiment for application in a low-income country. Health Policy Plann. 2009;24:151-8.
37. Coast J, Horrocks S. Developing attributes and levels for discrete choice experiments using qualitative methods. J Health Serv Res Policy. 2007;12:25-30.
38. De Bekker-Grob EW, Hol L, Donkers B, et al. Labeled versus unlabeled discrete choice experiments in health economics: an application to colorectal cancer screening. Value Health. 2010;13:315-23.
39. Rolfe J, Bennett J. The impact of offering two versus three alternatives in choice modelling experiments. Ecol Econ. 2009;68:1140-8.
40. Bridges JF, Buttorff C, Groothuis-Oudshoorn K. Estimating patients' preferences for medical devices: does the number of profile in choice experiments matter? National Bureau of Economic Research, 2011. Available from: http://www.nber.org/papers/w17482.
41. Kjaer T, Bech M, Gyrd-Hansen D,et al. Ordering effect and price sensitivity in discrete choice experiments: need we worry? Health Econ. 2006;15:1217-28.
42. Howard K, Salkeld G. Does attribute framing in discrete choice experiments influence willingness to pay? Results from a discrete choice experiment in screening for colorectal cancer. Value Health. 2009;12:354-63.
43. Veldwijk J, Lambooij MS, van Til JA, et al. Words or graphics to present a Discrete Choice Experiment: Does it matter? Patient Educ Couns. 2015;98:1376-84.
44. Orme BK. Getting started with Conjoint Analysis: Strategies for Product Design and Pricing Research. Madison, WI: Reserach Publishers LLC; 2006.
45. Mandeville KL, Lagarde M, Hanson K. The use of discrete choice experiments to inform health workforce policy: a systematic review. BMC Health Serv Res. 2014;14:367.
46. Vass C, Gray E, Payne K. Discrete choice experiments of pharmacy services: a systematic review. Int J Clin Pharm. 2015; [Epub ahead of print].
47. De Bekker-Grob EW, Ryan M, Gerard K. Discrete choice experiments in health economics: a review of the literature. Health Econ. 2012;21:145-72.

48. Ryan M, Gerard K. Using discrete choice experiments to value health care programmes: current practice and future research reflections. Appl Health Econ Health Policy. 2003;2:55-64.
49. Marshall D, Bridges JF, Hauber AB, et al. Conjoint Analysis Applications in Health - How are Studies being Designed and Reported?: An Update on Current Practice in the Published Literature between 2005 and 2008. Patient. 2010;3:249-56.
50. Crossing the quality chasm: a new health system for the 21st century. National Research Council. Washington, DC: National Academies Press; 2001.

PART I

The past - review of the literature

2

Discrete choice experiments in health economics: a review of the literature

Michael D. Clark
Domino Determann
Stavros Petrou
Domenico Moro
Esther W. de Bekker-Grob

Pharmacoeconomics. 2014;32(9):883-902
doi: 10.1007/s40273-014-0170-x

ABSTRACT

Background: Discrete choice experiments (DCEs) are increasingly used in health economics to address a wide range of health policy-related concerns.

Objective: Broadly adopting the methodology of an earlier systematic review of health-related DCEs, which covered the period 2001-2008, we report whether earlier trends continued during 2009-2012.

Methods: This paper systematically reviews health-related DCEs published between 2009 and 2012, using the same database as the earlier published review (PubMed) to obtain citations, and the same range of search terms.

Results: A total of 179 health-related DCEs for 2009-2012 met the inclusion criteria for the review. We found a continuing trend towards conducting DCEs across a broader range of countries. However, the trend towards including fewer attributes was reversed, whilst the trend towards interview-based DCEs reversed because of increased computer administration. The trend towards using more flexible econometric models, including mixed logit and latent class, has also continued. Reporting of monetary values has fallen compared with earlier periods, but the proportion of studies estimating trade-offs between health outcomes and experience factors, or valuing outcomes in terms of utility scores, has increased, although use of odds ratios and probabilities has declined. The reassuring trend towards the use of more flexible and appropriate DCE designs and econometric methods has been reinforced by the increased use of qualitative methods to inform DCE processes and results. However, qualitative research methods are being used less often to inform attribute selection, which may make DCEs more susceptible to omitted variable bias if the decision framework is not known prior to the research project.

Conclusions: The use of DCEs in healthcare continues to grow dramatically, as does the scope of applications across an expanding range of countries. There is increasing evidence that more sophisticated approaches to DCE design and analytical techniques are improving the quality of final outputs. That said, recent evidence that the use of qualitative methods to inform attribute selection has declined is of concern.

1. INTRODUCTION

Discrete choice experiments (DCEs) are increasingly used in health economics to address a wide range of health policy-related concerns. The approach draws its microeconomic foundations from the characteristics theory of demand [1] and random utility theory (RUT) [2]. The characteristics theory of demand assumes that goods, services, or types of healthcare provision can be valued in terms of their constituent characteristics (otherwise known as attributes). DCEs involve respondents making a number of stated preference choices in response to DCE questions. According to RUT, respondents are assumed to act in a utility maximizing manner and make choices contingent upon the levels of attributes in DCE scenarios. Therefore, choice data obtained from respondents' stated preferences can be analysed using econometric methods compatible with RUT. If the specified attributes are significantly related to respondent choices, findings from data analysis should confer information relating to how the average respondent's utility (or willingness to pay) is affected by changes in the levels of attributes. RUT assumes that respondent utility can be decomposed into a systematic component, which is a function of attributes and their levels, and a random component, which is an error term in the regression equation related to unmeasured preference variation. Published DCEs in healthcare are usually compatible with RUT [3-5], in the sense that they adopt a methodology consistent with RUT.

Although reviews and commentaries have been published of healthcare-related DCEs for specific clinical contexts or health-related concerns [6-8], and conjoint analyses more broadly [9], the most comprehensive reviews of the healthcare DCE literature cover the periods 1990-2000 [4] and 2001-2008 [3]. This paper updates those earlier systematic reviews to cover the period 2009-2012, and considers how key aspects of the design and application of DCEs have changed across the three periods.

2. METHODS

This review builds upon the earlier systematic reviews [3, 4]. It focuses on the 2009-2012 literature, and was derived from the literature in the sense that it replicates the methodology of the most recent review [3]. Although further checklists or commentaries on best practice [10-16] have been developed in recent years, these did not require fundamental changes to the approach to reviewing the DCE literature. Moreover, some of these checklists or commentaries [12-14] had already informed the development of the criteria deployed in the 2001-2008 review, whilst others served to confirm that our review encompassed appropriate criteria [11, 15, 16]. However, we did feel that the range of information extracted in relation to preference heterogeneity models such as mixed logit could be improved upon, and so we also gathered additional information on the distributional assumptions deployed when

mixed logit was applied, and the number of Halton draws that were specified for replications. Searches were restricted to the PubMed search engine, replicating the approach of the most recent review [3], and used the same search terms, including 'conjoint', 'conjoint analysis', 'conjoint measurement', 'conjoint studies', 'conjoint choice experiment', 'part-worth utilities', 'functional measurement', 'paired comparisons', 'pairwise choices', 'discrete choice experiment', 'dce', 'discrete choice mode(l)ling', 'discrete choice conjoint experiment', and 'stated preference'. Initial searches were conducted in September 2011, and then updated in March 2012 and August 2013 to ensure that all 2011 and 2012 papers were included. Moreover, for the period 2009-2012, we also allowed for the inclusion criteria to include a small number of best-worst scaling (BWS) DCEs/technical theoretical papers, as long as they generated health-related DCE results alongside BWS results. We included case 3 type BWS studies [17, 18] in the main review because, unlike case 1 and 2 BWS studies, these involve a comparison between two or more profiles [17]. This is despite the fact that case 3 BWS studies do differ from mainstream DCEs, because respondents choose the least attractive profile in addition to the most attractive one [17].

At the request of one of the peer reviewers, we also reviewed health-related adaptive conjoint analysis (ACA) and adaptive choice-based conjoint (ACBC) studies, and any menu-based conjoint analyses that had been published between 1990 and 2012. Such analyses had been excluded from the previous reviews [3, 4].

3. SEARCH RESULTS

Following the PubMed searches using the aforementioned search terms, we sourced 12 ACA/ACBC analyses for the period 2009-2012 [19-30]; 14 ACA/ACBC analyses for the period 2001-2008 [31-44], and no analyses of these types for the period 1990-2000. However, we reviewed these separately and deposited the data in an additional file because these analyses adopt a fundamentally different approach to valuing attributes to mainstream DCE analyses, and we wanted to ensure that data in the main body of the paper were in step with inclusion criteria adopted for the previous reviews [3, 4], which had excluded such analyses.

Overall, 179 analyses were identified as meeting the inclusion criteria for the main review, i.e. health-related DCEs or case 3 BWS analyses published in the English language. Each paper was read carefully and key data extracted systematically in the sense of evaluating them against a checklist of pre-established criteria, which largely corresponded to those used in the most recent of the previous reviews [3]. The data are summarized in the following Sects. 3.1-7.2. In Table 1, we provide further information relating to the definitions and other details of analyses in each of the categories.

Table 1 | Definitions and details of total number of analyses in each category.

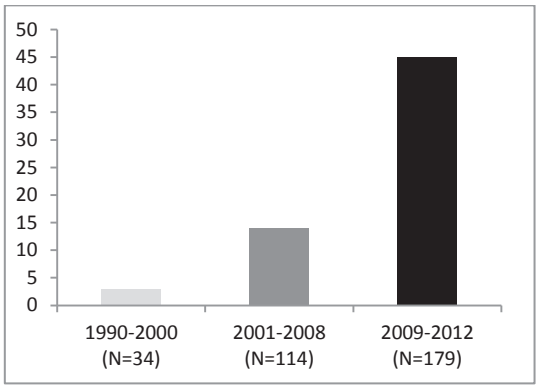
Category	Definition of category	Number of analyses (number of papers the analyses are contained in)	Number of papers per category also covered by other categories [citation details]	Other explanatory information
A	Patient or consumer experience factors	25 (24)	3 [47, 53, 62]	Within this category, a paper by Damman et al. [46] contains two analyses, one relating to knees and another relating to cataracts; an analysis by Goodall et al. [51] contains two analyses, one relating to patient preferences and another relating to carer preferences. There are also two papers cited for one analysis by Naik-Panvelkar et al. [60, 61]
B	Valuing health outcomes	13 (13)	2 [70, 72]	Not applicable
C	Investigating trade-offs between health outcomes and patient or consumer experience factors	81 (81)	13 [82, 90, 103, 107, 119, 133, 137, 139, 141, 143, 146, 158, 161]	Within this category there are two papers by Morton et al. [131, 132], which have been reviewed as one analysis. However, a paper by Regier et al. [143] contained two analyses (one relating to fungal treatment, and the other relating to bacterial issues)
D	Estimating utility weights within the quality-adjusted life-year framework	4 (4)	Not applicable	Not applicable
E	Job choices	11 (8)	Not applicable	Within this category, a paper by Rockers et al. [173] contains four analyses, one relating to medical student preferences, another relating to nursing student preferences, another relating to pharmacy student preferences, and another relating to science student preferences
F	Developing priority setting mechanism	24 (23)	8 [47, 71, 72, 82, 86, 107, 119, 161]	Within this category, the paper by Promberger et al. [184] contains two separate analyses for evaluation
G	Health professionals preferences for treatment and screening options	24 (23)	10 [53, 62, 70, 72, 103, 133, 139, 141, 143, 145]	Within this category, the paper by Regier et al. [143] contains two separate analyses for evaluation. One of these relates to fungal infection and the other relates to bacterial infection
H	Other	21 (21)	8 [63, 90, 103, 137, 145, 158, 159, 197]	Not applicable

Category A covers 25 separate analyses in 24 papers [45-67]; Category B includes 13 analyses in 13 papers [68-80]; Category C encompasses 81 analyses in 81 papers [81-162]; Category D relates to four analyses in four papers [163-166]; Category E covers 11 different analyses in eight papers [167-174]; Category F encompasses 24 analyses in 23 papers [47, 71, 72, 82, 86, 107, 119, 161, 175-189]; Category G includes 24 analyses in 23 papers [53, 62, 70, 72, 103, 122, 133, 139, 141, 143, 145, 190- 201]; and Category H relates to 'other' analyses, and there are 21 'other' analyses in 21 papers [63, 90, 103, 137, 145, 158, 159, 197, 202-214]. In Sects. 3.1-7.2, key findings are summarized in graphs. In the following text, we highlight changes that are of a reasonable magnitude and may be regarded as of significance.

3.1 Number of Discrete Choice Experiment (DCE) Analyses per Year

Figure 1 summarizes the average number of DCEs published per year across the three review periods. The 2001-2008 review [3] noted that the number of published applications of DCEs in healthcare rose from a mean of 3 per year (1990-2000) to a mean of 14 per year (2001-2008). Our review for 2009-2012 showed that the average number of analyses rose again to 45 per year (2009-2012), a marked increase, and peaked at 74 in 2012.

Figure 1 | Average number of DCE studies/year.

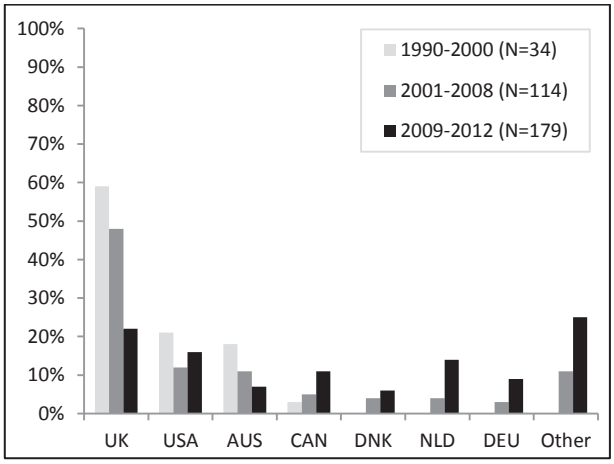


Abbreviation: *DCE* discrete choice experiment

3.2 DCE Studies Country of Origin

Figure 2 indicates the proportion of analyses emanating from different countries during the three time periods. The 2001-2008 review noted that the UK remained the main source of DCEs in healthcare. However, UK dominance has been eroded considerably since then (see Fig. 2). The proportion of analyses emanating from the UK has continued to fall, from 59% in 1990-2000 to 48% in 2001-2008, and to 22% during 2009-2012. Moreover, the proportion of analyses emanating from Australia (AUS) has fallen from 18% in 1990-2000 to 11% in 2001-2008, and to 7% in 2009-2012.

Figure 2 | Country of origin.



Abbreviations: *AUS* Australia; *CAN* Canada; *DNK* Denmark; *NLD* the Netherlands; *DEU* Germany

Comparing 2001-2008 with 2009-2012, an increased proportion of analyses now originate in the USA, Canada (CAN), Denmark (DNK), the Netherlands (NLD), and Germany (DEU). There was also an increase in analyses coming from 'other' countries (11% in 2001-2008 compared with 25% in 2009-2012), reflecting an increasing trend towards applying DCEs across a range of high-, middle-, and low-income countries[4].

3.3 The Number of Attributes Included in DCE Studies

Figure 3 provides information on the number of attributes included in DCE analyses across the three time periods. The most noteworthy change includes the fact that the proportion of analyses with four or five attributes rose from 29% in 1990-2000 to 44% in 2001-2008, but fell back to 32% in 2009-2012. There was also an increase in the number of DCEs with between seven and nine attributes: 12% in 1990-2000, 13% in 2001-2008, increasing to 22% in 2009-2012. The proportion of analyses with more than ten attributes fell from 12% in 1990-2000 to 2% in 2001-2008, and remained at that level in 2009-2012.

3.4 Domains of DCE Attributes

Figure 4 provides information on the proportion of DCEs encompassing different domains. The main noteworthy changes include the fact that the proportion of analyses with a domain related to time fluctuated; it was 74% in 1990-2000, in 2001-2008 it fell to 51%, before rising again to 65% in 2009-2012.

[4] Lower income countries in 2009-2012 included Kenya, South Africa, Thailand, China, Ghana, Vietnam, Ethiopia, Peru, Ukraine, India, Cuba, Nepal, Turkey, and Burkina Faso.

Figure 3 | Number of attributes.

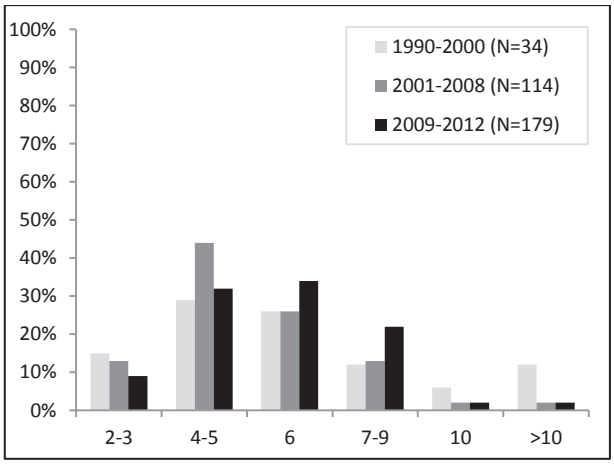
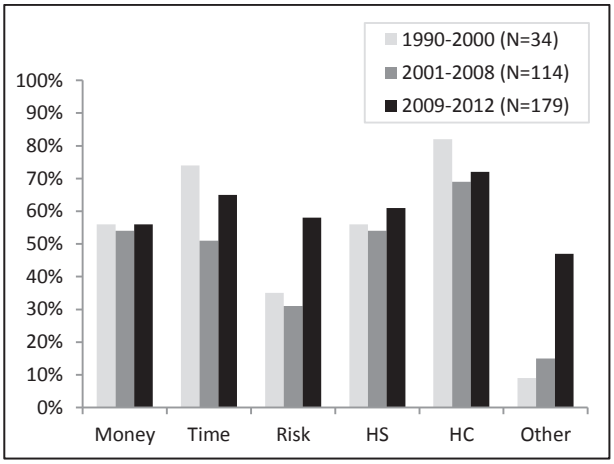


Figure 4 | Attribute domains.



Abbreviations: *HS* health status; *HC* healthcare

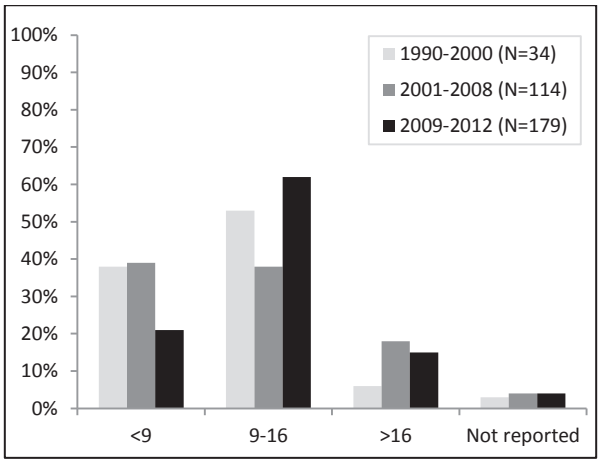
The proportion of DCEs including a measure of risk rose in the most recent period from 35% in 1990-2000 and 31% in 2001-2008, to 57% in 2009-2012.

The proportion of analyses with a healthcare (HC) domain also fluctuated: 82% in 1990-2000, falling to 69% during 2001-2008, and increasing to 72% in 2009-2012. At the same time, the proportion of analyses with attributes relating to 'other' domains, not encompassed by existing categories, increased from 9% in 1990-2000 to 15% in 2001-2008, and 47% in 2009-2012. Additional categories in Fig. 4 include a monetary domain (Money) and a health status domain (HS).

3.5 The Number of Questions Posed by DCEs

Figure 5 provides information on the number of choice tasks posed by DCEs. The main noteworthy trends are as follows. The proportion of DCE analyses posing eight or fewer choices (<9) was 38% in 1990-2000, 39% in 2001-2008, but fell back to 22% in 2009-2012. In contrast, the proportion of analyses with 9-16 choices was 53% in 1990-2000, falling to 38% during 2001-2008, and rising to 62% during 2009-2012. The proportion of analyses with more than 16 choices (>16) rose initially and then stabilized (6% in 1990-2000, 18% in 2001-2008, and 15% in 2009-2012).

Figure 5 | Number of choice tasks.



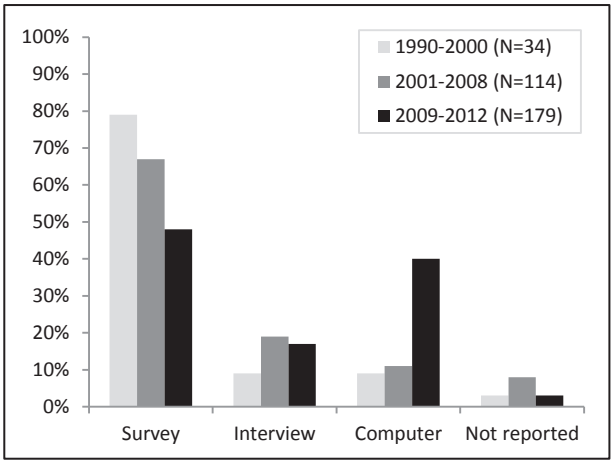
3.6 DCE Survey Administration

Figure 6 provides information on the different modes of administering DCEs. Since 1990, there has been a trend away from self-completed pen/paper questionnaires. The proportion of analyses using self-completed questionnaires (Survey) was 79% in 1990-2000, falling to 67% in 2001-2008, and then further to 48% in 2009-2012. The proportion of interviewer-administered DCEs (Interview) was 9% in 1990-2000, rising to 19% in 2001-2008, and was 17% in 2009-2012.

Overall, there has been a trend towards DCEs involving computerized administration (Computer), sometimes involving internet surveys, over the last 20 years. There have been improvements in computer technology, combined with the increased use of computers by the wider population. This has made accessing DCE respondent samples using computers easier. Moreover, the ease with which DCE samples can be accessed using computers and the internet partly explains the trend towards increased use of DCEs since 1990. During 1990-2000, 9% of analyses involved computerized administration; the figure was 11% in 2001-2008 and then rose sharply to 40% during 2009-2012. As depicted in Fig. 6, a small

proportion of analyses failed to report (Not reported) the form of survey administration in each period.

Figure 6 | Survey administration.



4. DCE EXPERIMENTAL DESIGN AND CHOICE SET CONSTRUCTION

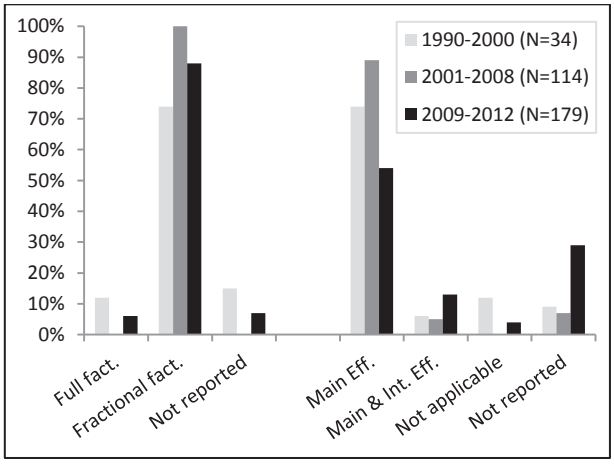
A good discussion of some of the relevant issues relating to DCE design is contained in the review by de Bekker-Grob et al. [3], so in the interests of brevity, we refer the reader to that paper and to another key citation [11].

4.1 Design Plan

Figure 7 depicts information on the different types of design plans. The most noteworthy trends are as follows. The proportion of analyses involving full factorial (Full fact.) DCE designs fell from 12% in 1990-2000 to 0% in 2001-2008, and then rose again to 6% during 2009-2012. In the period 1990-2000, 74% of analyses adopted a fractional factorial design (Fractional fact.), a proportion that rose to 100% in 2001-2008, but then decreased to 88% in 2009-2012. Overall, 15% of analyses did not clearly report their design type (Not reported) in 1990-2000, a proportion that fell to 0% in 2001-2008, but then rose again to 7% during 2009-2012.

Overall in 1990-2000, 74% of DCE analyses involved a 'main effects' design (Main Eff.), and this proportion rose to 89% during 2001-2008, but then fell back to 54% in 2009-2012. Therefore, as with the baseline and 2001-2008 reviews, 'main effects' designs remain the dominant type of design in published DCE studies. In 1990-2000, 6% of analyses catered for interaction effects alongside main effects (Main & Int. Eff.); the proportion was 5% in 2001-2008, and increased to 13% in 2009-2012. In some cases, a design plan was not applicable, whilst in others, it was not reported.

Figure 7 | Design plan.



Abbreviations: *fact*. factorial; *Eff*. effects; *Int*. interaction

4.2 Use of Software Packages to Design DCEs

Figure 8 summarizes information on the use of different software packages for the design of the DCEs. The most noteworthy trends are that the use of a software package to design DCEs remained steady throughout the years (Fig. 9).

Figure 8 | Software packages.

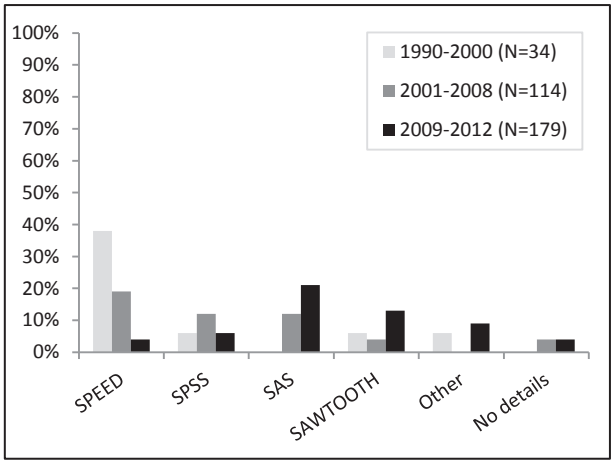


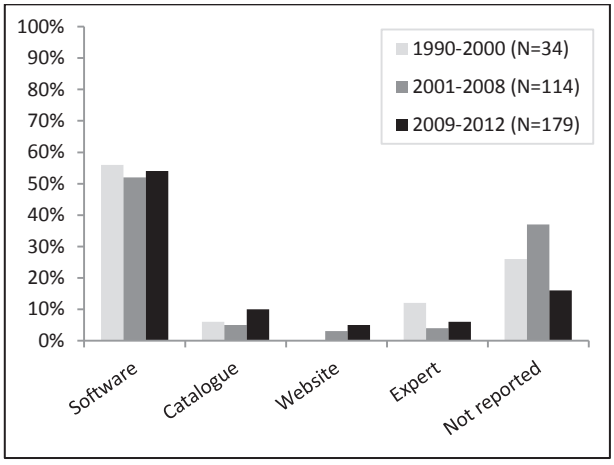
Figure 8 shows there seems to have been a general trend away from using SPEED over the period as new software has become available. In 1990-2000, 38% of analyses used SPEED; this proportion fell to 19% in 2001-2008, and to 4% in 2009-2012.

The SAS package (which can provide D-efficient designs) has become increasingly popular. Recorded use rose from 0% in 1990-2000 to 12% in 2001-2008, and 21% in 2009-2012. Use of SPSS did not change that much. The use of Sawtooth software fluctuated; 6% of analyses used this software in 1990-2000, 4% used it in 2001-2008, and 13% used it in 2009-2012. The use of 'other' software was also low (6% in 1990-2000, 0% in 2001-2008, and 7% in 2009-2012[5]), and only a small proportion of analyses in each period failed to provide information on type of software (No details).

4.3 Use of Design Catalogues, Websites, and Expert Advice to Design DCE Questionnaires

Figure 9 depicts information on the use of software, design catalogues, websites, and experts to inform DCE design and whether this was not reported. Figure 9 shows that there have not been any particularly large changes in the use of these over the three periods.

Figure 9 | Design source.

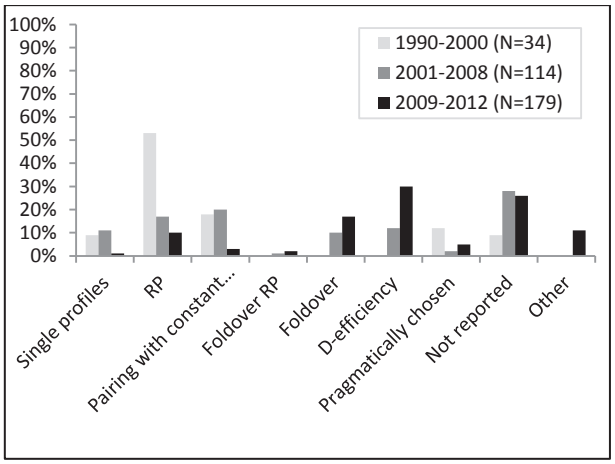


4.4 Methods Used to Create Choice Sets

Figure 10 depicts information on the use of different methods to create choice sets. The most noteworthy trends are as follows. In 1990-2000, 9% of analyses reported designs that involved orthogonal arrays with single profiles, i.e. binary choices (Single profiles); the proportion was 11% in 2001-2008, but fell to 1% in 2009-2012. Use of orthogonal arrays with random pairing (RP) was more common, but has fallen over time; in 1990-2000 it was applied in 53% of analyses, falling to 17% in 2001-2008 and 10% in 2009-2012.

[5] 'Other' packages used included Gauss for two analyses; Ngene (a Bayesian efficient design) for four analyses; and the statistical design procedure Gosset for one analysis; a D-efficient design advocated by Rose and Bliemer for one analysis; STATA for one design; a design described as "an experimental design algorithm optimizing orthogonality, attribute balance, and efficiency" for one design; and Street and Burgess Software for one design.

Figure 10 | Methods to create choice sets.



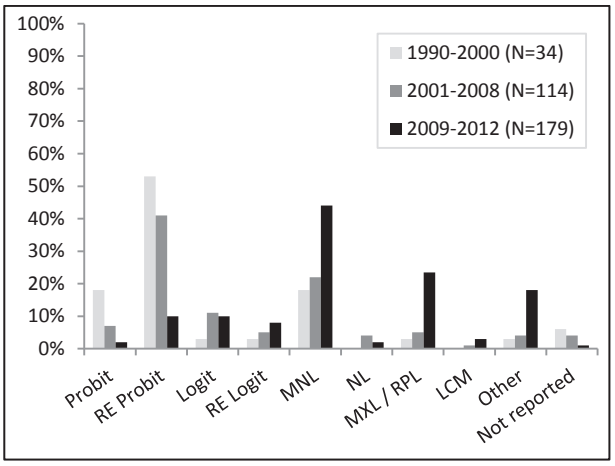
Abbreviations: *RP* random pairing; *constant* constant comparator

Analyses involving orthogonal arrays with pairing with a constant comparator constituted approximately one in five designs in earlier periods, 18% in 1990-2000 and 20% in 2001-2008, before falling to 3% in 2009-2012. In 1990-2000, none of the analyses involved orthogonal arrays with a foldover design, but this proportion rose to 10% in 2001-2008 and 17% in 2009-2012. Very few analyses in each period used a foldover design with random pairing (Foldover RP), or pragmatically chosen designs. Similarly, there has been a general trend towards D-efficient designs (D-efficiency), rising from 0% in 1990-2000 to 12% of studies in 2001-2008, and 30% of studies in 2009-2012. The proportion of analyses that did not clearly report (Not reported) the methods used to create choice sets rose from 9% in 1990-2000 to 28% in 2001-2008 and stabilized at 26% in 2009-2012, whilst in one period (2009-2012), 11% of analyses used 'other' methods to create choice sets.

5. ESTIMATION PROCEDURES

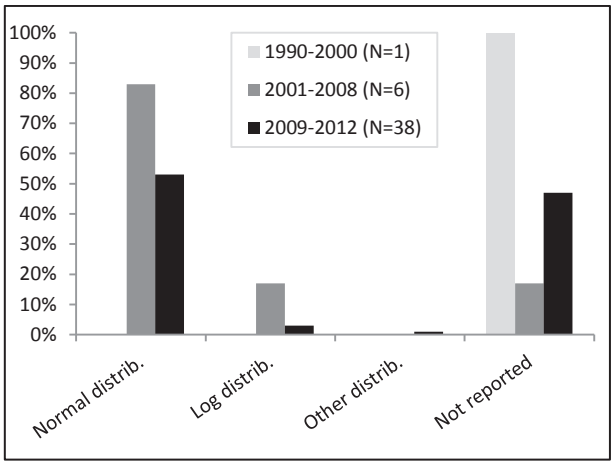
As there is a good explanation of RUT, alternative DCE econometric models, and the appropriateness of different models for different DCE applications in the earlier review paper, we refer readers to Sects. 5.1–5.3 of that paper in the interests of brevity [3]. Figure 11 depicts information on the different estimation methods. The most noteworthy trends are described in the sections below. In Fig. 11 details of the econometric estimation methods used are depicted.

Figure 11a | Estimation procedures.



Abbreviations: *RE* random effects*; MNL* multinomial logit; *NL* nested logit; *MXL / RPL* mixed logit / random parameters logit; *LCM* latent class model

Figure 11b | Distributional assumptions.



Abbreviation: *Distrib*. distributions

5.1 Use of Probit, Random Effects Probit, Logit, and Random Effects Logit

As previously reported, early DCE studies, i.e. those published in 1990-2000, seemed to focus upon applying either binary choice or 'forced choice' DCEs [4]. So, for example, in 1990-2000, 18% of analyses used probit; this proportion fell to 7% in 2001-2008, and fell further to 2% in 2009-2012. Similarly, in 1990-2000, 53% of analyses used random effects probit (RE probit), falling to 41% in 2001-2008, and then further to 10% in 2009-2012. The proportion of logit analyses was 3% in 1990-2000, rising to 11% in 2001-2008, and was 10% in 2009-2012, and relatively few analyses used random effects logit (RE Logit) in each time period.

5.2 Use of Multinomial Logit

The overall decline in the use of logit, probit and random effects probit reported above has been offset by an increased use of multinomial logit (MNL) analyses, which are sometimes otherwise known as conditional logit analyses. These analyses have the advantage that they can cater for more than two response options, and they can also allow respondents to 'opt-out' from making a decision. Sometimes such models may also be associated with better 'goodness of fit' than some other econometric models. During 1990-2000, 18% of studies used MNL, 22% used it during 2001-2008, rising to 44% during 2009-2012.

5.3 Use of Nested Logit

During the period 2001-2008, a small shift towards use of nested logit (NL), a technique that relaxes the independence of irrelevant alternatives (IIA), was observed [3]. It was applied in 4% of studies during 2001-2008, up from 0% in 1990-2000. For the period 2009-2012, the proportion remained low at 2% of studies.

5.4 Models Applicable When There is Preference Heterogeneity

During 1990-2000, only 3% of studies used mixed logit/random parameters logit (MXL/RPL); by 2001-2008, 5% of analyses used MXL/RPL. During the period 2009-2012, there was a clear trend towards increased use of MXL/RPL, and 21% of analyses utilized the technique. All the analyses involving MXL/RPL found evidence of preference heterogeneity. Ideally, when MXL/RPL analyses are submitted for publication, details of the number of replications (sometimes described as 'Halton draws') should be provided, as results can be sensitive to the number of replications. This occurred in 0% of MXL/RPL analyses in 1990-2000, 67% in 2001-2008, and 47% in 2009-2012.

Unlike latent class models (LCM), MXL/RPL analyses make distributional assumptions for random parameters (Fig. 11b). Not all MXL/RPL analyses indicate what these are. Indeed, in 1990-2000, 100% of analyses failed to indicate them. In 2001-2008, 17% failed to provide this information, and in 2009-2012, 47% failed to provide this information. When such information was provided, analyses usually indicated they had assumed normal distributions for random parameters: 0% in 1990-2000, 83% in 2001-2008, and 53% in 2009-2012. However, in 17% of cases in 2001-2008, and 8% in 2009-2012, models assuming logarithmic distributions for random parameters were also reported alongside results from models assuming normal distributions for random parameters. Also, one study [115] used a mixed logit hierarchical Bayesian model (MLHB), an extension of mixed logit modelling. Another analysis used what it described as a Bayesian-like approach similar to mixed logit [80]. Sometimes hierarchical Bayesian analysis has also been used without mixed logit [48, 70, 114, 204]. A small proportion of analyses did not report the estimation procedure used.

During the early period (1990-2000), no study used LCM. During the period 2001-2008, one study (1%) used LCM, and the study also found evidence of preference heterogeneity [216]. During 2009-2012, five analyses (3%) [57, 61, 129, 168, 204] used LCM, and they all identified evidence of preference heterogeneity.

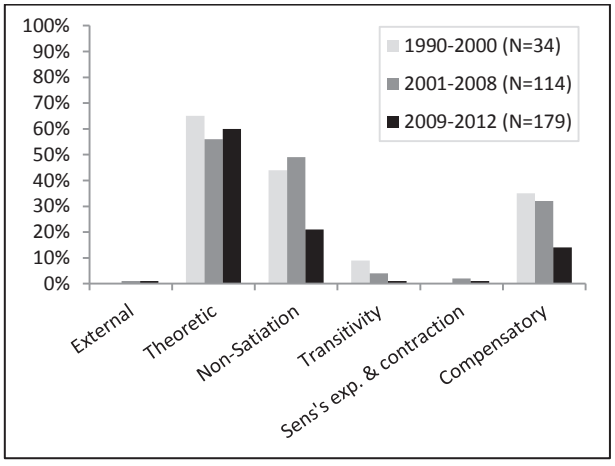
A few analyses used 'other' estimation procedures. In 1990-2000, this was the case for 3% of analyses; the proportion was 4% of analyses in 2001-2008 and 17% of analyses in 2009-2012[6].

6. VALIDITY

6.1 Validity Checks

The proportions of studies that used different validity tests are depicted in Fig. 12.

Figure 12 | Validity checks.



Abbreviation: *Exp.* Expansion

[6] 'Other' methods used in 2009-2012 included weighted probit [68]; OLS with a hetero-robust covariance matrix estimator [192]; a method described as ''modelling including interaction effects'' [45]; Cox's proportional hazards model with time-dependent covariate [105]; weighted least squares regression to estimate utility weights [105]; multivariate ordered probit to estimate conjoint utility parameters [76]; mixed logit with hierarchical Bayesian modeling and ordered probit [115]; generalized estimated equations [109, 125]; random parameter logit estimated using a hierarchical Bayesian algorithm [208]; conditional logit and parameter weighting functions [160]; a series of multivariate regressions [50, 65]; a method described as Bayesian-like for preference weights [80]; OLS [87]; hierarchical Bayesian analysis [48, 70, 114, 205, 212]; multinomial exploded logit [177]; Firth's unbiased estimator [193]; combined conditional logit and ranked logit model [127]; multivariate multilevel logistic regression [46]; generalized multinomial logit [119]; mixed effect logistic regression [184], error components mixed logit analysis [63]; a combination of Bayes theorem, Monte Carlo Markov chain procedure and the Metropolis Hastings algorithm [182]; and logistic and probit regression using cluster-robust standard error (SE), random effects and GEE and multinomial logistic and probit regressions with cluster-robust SE and random effects multinomial logistic model and probit model with cluster-robust SE treating the choices from two stages as two correlated binary outcomes [94].

The most noteworthy trends are as follows. Tests of external validity (External) are particularly valuable because stated preferences from DCEs can then be compared with revealed preferences. However, there is often little scope to conduct tests of external validity (particularly if DCEs are applied in the context of state funded healthcare provision). This may explain why none of the analyses contained a test for external validity during 1990-2000. The proportion rose to 1/114 analyses (1%) in 2001-2008 as there was a study [217] that compared doctors' prescribing decisions in relation to prescriptions for alcoholism with the preferences they expressed in a DCE. For the most recent period (2009-2011), only one [144] analysis (<1%) contained a test of external validity.

Most analyses included tests for internal theoretical validity (Theoretic). Overall, 65% of analyses in 1990-2000 included these tests, with the proportion falling to 56% in 2001-2008, and it was 60% in 2009-2012. Such tests involve an assessment of whether coefficients appear to move in line with prior expectations, and studies generally reported that this was the case.

Tests for non-satiation were less frequently reported. For the period 1990-2000, 44% of analyses contained such a test; the proportion was 49% in 2001-2008, before falling to 21% in 2009-2012. The decline in the use of such tests probably reflects concerns that they tend to be passed, so that they are a relatively weak test of validity. Also, in an influential paper [218], it has been argued that preferences that may appear to be 'irrational' may in reality be compatible with some form of rationality. Therefore, to delete such responses may be inappropriate, so the decline in the use of such tests may reflect good practice.

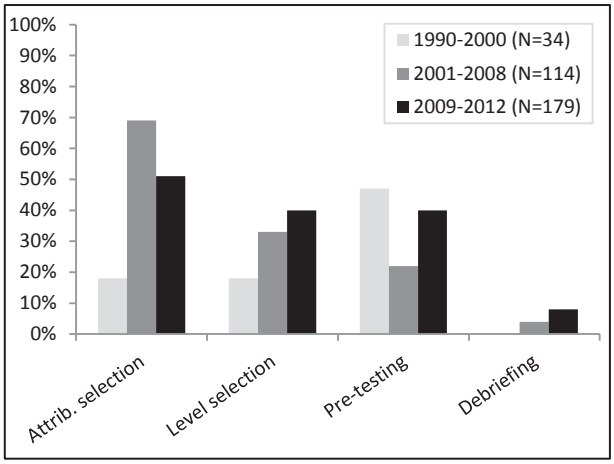
If tests of transitivity could readily be applied using DCEs, the information yielded might be more useful. However, they cannot always be readily applied, which is presumably why over the period 1990-2000 only 9% of analyses contained a transitivity test; in 2001-2008, 4% of analyses contained such a test, and during 2009-2012, 1% of analyses contained a transitivity test. During 1990-2000, none of the analyses contained a test relating to Sen's expansion and contraction properties (Sen's exp. & contraction); the proportion was 2% of analyses during 2001-2008, and 1% of analyses during 2009-2012. Use of a test for internal compensatory decision-making (Compensatory) [3] was much more frequent. In 1990-2000, 35% of analyses involved such a test; the value for 2001-2008 was 32% of analyses, but in 2009-2012, this declined to 14% of analyses.

6.2 Use of Qualitative Methods to Enhance DCE Processes and Results

Information on the use of qualitative methods to enhance DCE processes and results is depicted in Fig. 13. In 1990-2000, 18% of analyses used qualitative methods to inform attribute selection (Attrib. selection), rising to 69% of analyses in 2001-2008, before declining to 51% of analyses in 2009-2012. This is potentially a worrying trend because if the

selection of attributes is not properly grounded in qualitative research, then inappropriate attributes may be specified and appropriate attributes may be omitted (triggering omitted variable bias). It would be of little concern, however, if the recent reduction in the use of qualitative methods to inform attribute selection was triggered by the wider use of DCEs in contexts in which the decision framework is already known (for example, if DCEs are conducted alongside clinical trials).

Figure 13 | Use of qualitative methods.



Abbreviation: *Attrib*. attribute

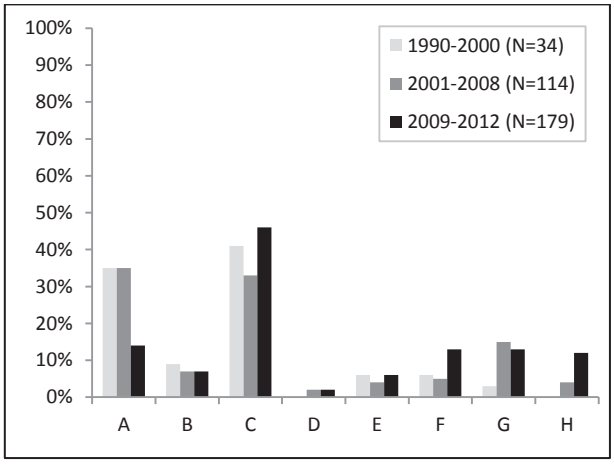
In contrast, the use of qualitative methods to inform attribute level selection (Level selection) increased; the proportion was 18% of analyses in 1990-2000, increasing to 33% in 2001-2008, before increasing again to 40% in 2009-2012. The use of a pre-testing questionnaire (Pre-testing) fluctuated over time; it was applied in 47% of analyses in 1990-2000, just 32% of analyses in 2001-2008, but was applied in 41% of analyses in 2009-2012. The use of debriefing choices (Debriefing) to help strengthen understanding increased from 0% of analyses in 1990-2000 to 4% of analyses in 2001-2008, and 8% of analyses during 2009-2012.

7. AREAS OF APPLICATION AND OUTCOME MEASURES

7.1 Areas of Application

As indicated by de Bekker-Grob et al. [3], although DCEs had originally been introduced into health economics primarily in order to value patient experience [219], there was clear evidence that the application of DCEs had broadened considerably by 2000-2008 [3]. Moreover, this trend continued into 2009-2012. Figure 14 summarizes the relevant data (for the definitions of categories A, B, C, D, E, F, G, and H, refer to Table 1).

Figure 14 | Areas of application.



Notes: For the definitions of categories A, B, C, D, E, F, G, and H, refer to Table 1.

The main noteworthy trends are as follows. In 1990-2000, 35% of analyses had a main study objective that involved valuing experience factors (Category A). The proportion was the same in 2001-2008, as 35% of analyses had the same main study objective. However, during 2009-2012, this proportion fell to 12% of analyses. In contrast, the proportion of DCEs exploring trade-offs between health outcomes and experience factors has risen steadily (Category C). In 1990-2000, 41% of analyses had this as a primary objective; during 2001-2008, the proportion was 33% of analyses, increasing further to 44% of analyses in 2009-2012. This reflects a shift from examining patient experience factors (Category A) in isolation (down from 35% in 2001-2008 to 12% in 2009-2012) towards estimating trade-offs between health outcomes and experience factors (Category C), which increased from 33% of analyses in 2001-2008 to 41% in 2009-2012. The latter category (Category C) also includes the estimation of trade-offs for non-patient groups, whereas the former (Category A) is specific to patient respondents.

In 1990-2000, no analysis included a main objective of estimating utility weights within a quality-adjusted life year (QALY) framework (Category D). During 2001-2008, 2% of analyses had this as the main objective. These two analyses used DCEs as an alternative to standard gamble (SG) and time trade-off (TTO) to estimate utility weights within a QALY framework [220, 221]. More recently, there has been some further work in this area. In 2009-2012, 2% of analyses had this as their main objective, reflecting some interest in this research agenda. One of these analyses [165] looked at deriving distributional weights for QALYs using DCEs; another [164] used DCEs to quantify EQ-5D health states, whilst another [163] explored whether a DCE that resembles TTO exercises is able to estimate consistent values on the health utility scale for the EQ-5D. A further analysis compared case 3 BWS DCE analysis with WTP analysis [166].

A small proportion of DCEs have had a primary objective of evaluating job choices, human resource policy (Category E), or developing priority setting frameworks (Category F), and values for these categories did not exhibit much change (see Fig. 14).

During 1990-2000, 3% of analyses had the main objective of establishing health professionals' preferences (Category G); this rose to 15% in 2001-2008, before falling slightly to 12% in 2009-2012. In 1990-2000, 0/34 analyses (0%) had an 'other' (Category H) main objective; this rose to 4% of analyses in 2001-2008, before rising again to 10% of analyses in 2009-2012[7].

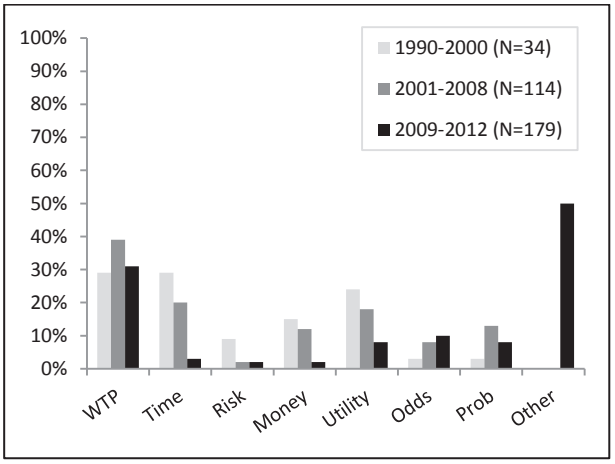
7.2 Outcome Measures

Information on trends relating to 'main outputs' is depicted in Fig. 15. In the past, DCEs often expressed outputs in terms of a primary outcome measure of 'per WTP unit' or 'per time unit'. In 1990-2000, 29% of analyses used the 'per WTP unit' (WTP) outcome measure, increasing to 39% of analyses in 2001-2008; however, in 2009-2012, the proportion was only 31% of analyses. The use of 'per unit of time' (Time) as an outcome measure has also declined markedly over the period. During 1990-2000, 29% of analyses used this outcome measure; in 2001-2008, the proportion was 20% of analyses, and it declined further to 3% of analyses in 2009-2012.

The proportion of DCEs using 'per risk unit' (Risk) as a primary outcome is low, and this has fluctuated little over the period (see Fig. 15). Only a minority of analyses use monetary welfare measures (Money) as the primary outcome measure, and this proportion has fallen in proportionate terms over time. During 1990-2000, 15% of analyses involved a money welfare measure; during 2001-2008, the proportion was 12% of analyses; and during 2009-2012, the proportion was 2% of analyses.

7 In 2009-2012, one study explored how changing the number of responses elicited from respondents might affect estimates of WTP [204]; another looked at parents' preferences for management of attention-deficit hyperactivity disorder [206]; one study looked at general public preferences for long-term care [137]; another two studies looked at preferences for human papillomavirus vaccine, one case looking a societal preferences [207] and the other [63] looking at mothers' preferences; another study looked at the valuation of diagnostic testing for idiopathic developmental disability by the general population [208]; another looked at various stakeholder groups' preferences for coagulation factor concentrates to treat hemophilia [145]; one study looked at general public preferences for tele-endoscopy services [158]; another compared Dutch and German preferences for health insurance amongst their populations [214]; one paper looked at public and decision maker preferences for pharmaceutical subsidy decisions [215]; one study explored how individuals perceive various coronary heart disease factors [203], whilst another described the relative importance of major adverse cardiac and cerebrovascular events to be used when analyzing trials [212]. Two other DCEs were performed on the area of quality improvement; one investigated how to best disseminate evidence-based practices to addiction service providers and administrators [205], while the other was used to investigate which indicators had the greatest impact on the decisions of health service inspectors concerning the assessment of quality of mental health care [211]. Other applications included a study on preferences of health workers in Burkina Faso for health insurance payment mechanisms [209]; a study on how respondents valued mortality risk attributable to climate change reductions [210]; and a study on the preferences for reducing contaminated sites to reduce the risk for cancer [213].

Figure 15 | Outcome measures.



Abbreviations: *WTP* willingness to pay; *Prob*. Probability

The use of utility scores (Utility) as the primary outcome measure is more common, and this has fluctuated over time. In 1990-2000, 24% of analyses had utility scores as the primary outcome measure. The proportion was 18% of analyses during 2001-2008, decreasing to 8% of analyses during 2009-2012.

There is also evidence that the use of odds ratios (Odds) has fluctuated. In 1990-2000, only 3% of analyses used odds ratios as the primary outcome measure. By 2001-2008, the proportion of had more than tripled to 11% of analyses; the proportion was 8% in 2009-2012. Likewise, the use of probability scores (Prob) increased from 3% of analyses in 1990-2000 to 13% of analyses in 2001-2008, before declining to 8% of analyses during 2009-2012.

Finally, Fig. 15 presents information on 'other' outcome measures used. For the periods 1990-2000 and 2001-2008, the earlier review authors did not use an 'other' category for the main outcome measure used. However, for this review, we categorized a substantive number of analyses, 49% of analyses, in the 'other' category. This was mainly because in 2009-2012 there was a trend to use importance scores or relative importance of attributes (25% of analyses) or preference weights (6% of analyses).

8. REVIEW OF ADAPTIVE CONJOINT ANALYSIS/ADAPTIVE CHOICE BASED CONJOINT STUDIES

Having summarized the ACA/ACBC studies in an additional file (because they use a fundamentally different approach to DCEs), we concluded that it was difficult to discern major trends in relation to these analyses across study periods. This is because our PubMed

search unearthed a total of only 26 analyses in 26 papers to review for the entire review period, 1990-2012. Nonetheless, for the interesting trends that have been discerned, the details have been provided in additional file 1.

Our searches (Table A1; see additional file 1 for all appendix Tables) unearthed no studies of this type for the period 1990-2000, 14 analyses for the period 2001-2008, and 12 analyses for the period 2009-2012. In contrast to the DCE literature, most of these analyses seemed to emanate from the USA, and the surveys were more likely to be computer administered. The analyses also tended to be designed using Sawtooth software (Table A2), and no analyses indicated the use of an alternative software package. Published analyses did not indicate whether they ever used the main estimation methods used for DCEs (Table A3), and all of them either fell into the 'other' or 'not reported' categories. Moreover, validity checks were rarely incorporated into ACA/ACBC analyses (Table A4).

9. DISCUSSION

The number of published health-related DCEs has increased dramatically over the last 2 decades. There has also been a shift away from UK dominance of health-related DCEs, with a widening range of countries producing DCE studies.

The 2001-2008 review [3] noted a wide range of policy applications for DCEs; this continued during 2009-2012. In 2001-2008, the valuation of patient experience continued ''to be the focus of the majority of studies'' [3]. In contrast, this declined as the main focus during 2009-2012. Nevertheless, in 2009-2012, most DCEs continued to include attributes relating to patient experiences, but increasingly in order to examine trade-offs between health outcomes and experience factors. Also health outcomes and experience factors for respondents groups other than patients are encompassed by this category (e.g., Category C analyses).

The 2001-2008 review [3] reported that ''willingness to pay continued to be a commonly used output from DCEs'' over that period. However, the present review found evidence that the proportion of DCE studies using either a 'per WTP unit' or a 'monetary welfare measure' as their primary outcome has fallen. This could in part be attributable to concerns that have been raised in relation to the use of DCEs to elicit WTP. These include whether estimated WTP obtained via DCEs may be sensitive to the range specified for the monetary attribute or the presence or absence of payment per se [222], or the presence or absence of a non-zero cost, rather than the level of cost indicated by the monetary attribute [223]. Other evidence suggests that the way attributes are 'framed' in a DCE questionnaire may impact upon estimated WTP [113]. Furthermore, the hypothetical nature of DCEs can hinder correct estimates of WTP because respondents will not be bound by the choices they make [143].

Another issue arises because when estimating marginal willingness to pay (MWTP), it is commonly assumed that marginal utility of money is constant and the cost function is therefore continuous and linear. However, there is reason to believe that the cost gradient may not be continuous and linear [115]. Therefore, if WTP is calculated there is a need to proceed with econometric analysis that assumes more complex indirect utility functions [123, 124], for example, using interaction terms between attributes [224], or using non-linear attribute transformations, such as the squaring of attributes [225], or taking natural logarithms [226].

In addition to using methods that can be used to identify unobserved preference heterogeneity, which we discussed in detail in Sect. 5.4, the issue of segmenting DCE data to examine the preferences of defined subgroups is sometimes important. One early analysis [227] segmented the DCE data according to the severity of symptoms associated with osteoarthritis, and the importance of a joint ache attribute was seen to increase in respondents with more severe symptoms. Other analyses relating to establishing priority criteria for allocating cadaveric kidney transplants [71, 72] have provided evidence of statistically significant differences in preferences between different stakeholder groups. A major finding to emerge from this research was that whilst non-ethnic minority patients would prefer to allocate kidney transplants to recipients with a good tissue match, ethnic minorities (who would be disadvantaged by use of such priority criteria) would not. Another interesting analysis relating to segmentation used segmentation because ''health organizations need to understand whether the same health treatments, prevention programs, services, and products should be applied to everyone in the relevant population or whether different treatments need to be provided to each of several segments that are relatively homogeneous internally but heterogeneous among segments'' [228]. Segmenting the data to facilitate subgroup analysis is particularly appropriate if policy-relevant differences in preferences between defined subgroups might be applicable.

The use of simulation may be important, when DCEs are applied, because simulation may enable you to do something useful with DCE data. For example, in a DCE relating to junior doctors' preferences for specialty choice, it was found that increasing general practitioners (GPs) wages by $AUS50,000, or increasing opportunities for procedural or academic work, can increase the number of junior doctors choosing general practice by between 8 and 13% [174]. Another example of how useful simulation of DCE data can be is an analysis that was designed to predict the place of out-of-hours care. Using DCE data to predict market shares, it was predicted that a new GP cooperative could capture about a third of the market, ahead of the emergency department, the second most preferred service [183].

In Sect. 5.2, we pointed out that econometric methods are increasingly being used which can facilitate allowing respondents to 'opt-out' from registering a preference. An example

of why this might be important is apparent from an analysis relating to colorectal cancer screening [110]. This is because when screening for colorectal cancer, it was important to give people an 'opt-out' response in order to ensure that the choices respondents faced were realistic. Similarly, when evaluating two hypothetical smoking cessation mechanisms, it was important to provide respondents with an 'opt-out' option [128].

This review for 2009-2012 found an increasing trend towards presenting respondents with more DCE choice scenarios. Some evidence suggests [203] that later DCE responses might be more thought through, so this may be a welcome development. However, the optimal number of choices presented should ideally be established through piloting, because it is likely to be a function of the complexity of choices presented. There has also been an interest in developing approaches to cater for the inclusion of increased numbers of attributes within DCE designs [161, 198]. Another development has been a shift away from self-administered pen/pencil-response DCE questionnaires towards either interviewer-administered or, more particularly, computer-administered questionnaires and the use of internet panels.

The trend towards the increased use of D-efficient DCE designs noted by de Bekker-Grob et al. [3] has continued. Although there has been an increase in the proportion of analyses catering for interactions, main effects designs remain dominant. Trends away from the use of probit and random effects probit towards greater use of MNL reflect the increased use of DCEs incorporating more than two choices, or two choices plus an opt-out. Recent interest in the use of models catering for preference heterogeneity is welcome because when mixed logit or LCM is applied, it invariably identifies preference heterogeneity, which otherwise would have been overlooked. That said, one limitation of the mixed logit approach is that it requires the imposition of assumptions about the distribution of the random coefficients. If the distributional assumptions are not valid, then this can undermine the validity of findings about preference heterogeneity. LCM has the advantage over mixed logit that it does not involve the imposition of such distributional assumptions, but can have the disadvantage that it may be more time consuming to implement.

There appears to have been a decline in the use of most validity tests during 2009-2012, including tests of non-satiation, transitivity, Sen's expansion and contraction properties, and tests of compensatory decision making. We might have anticipated a decline in the use of such validity checks, because the usefulness of the results they yield has increasingly been called into question [144, 218, 229]. Reassuringly, however, during 2009-2012, there has been an improvement in the proportion of analyses using qualitative methods to enhance DCE processes and results, in some respects, including the use of qualitative methods to inform level selection; use of pilot pre-testing questionnaires; and the use of qualitative methods to strengthen understanding of responses (including debriefing choices). A remaining cause for concern, however, is that the use of qualitative methods to inform

attribute selection has declined since 2001-2008. This could lead to increased likelihood of omitted variable bias affecting DCE results.

The main limitation of this study was that, like the published review for 2001-2008 [3], we only used PubMed to source literature. However, as that review noted [3], when additional searches are conducted on other databases, it does not markedly affect review findings. So in the interests of ensuring comparability with data from that earlier review, we also restricted our searches to the PubMed database.

10. CONCLUSIONS

The use of DCEs in healthcare continues to grow dramatically, as does the scope of applications across an expanding range of countries. There is increasing evidence that more sophisticated approaches to DCE design and analytical techniques are improving the quality of final outputs. That said, recent evidence that the use of qualitative methods to inform attribute selection has declined is of concern.

REFERENCES

1. Lancaster K. New approach to consumer theory. J Polit Econ. 1966;74:132–57.
2. McFadden D. Computing willingness-to-pay in random utility models. Trade theory and econometrics, chap. 15. In: Essays in honour of John S. Chipman. Studies in the Modern World Economy; 1999. p. 253–74.
3. De Bekker-Grob EW, Ryan M, Gerard K. Discrete choice experiments in health economics: a review of the literature. Health Econ. 2012;21:145–72.
4. Ryan M, Gerard K. Using discrete choice experiments to value health care programmes: current practice and future research reflections. Appl Health Econ Health Policy. 2003;2:55–64.
5. De Bekker-Grob EW, Chorus CG. Random regret-based discrete-choice modelling: an application to healthcare. Pharmacoeconomics. 2013;31:623–34.
6. Petrou S, McIntosh E. Commentary: Using stated preference discrete choice experiments to elicit women's preferences for aspects of maternity care. Birth. 2011;38:47–8.
7. Lagarde M, Blaauw D. A review of the application and contribution of discrete choice experiments to inform human resources policy interventions. Hum Resour Health. 2009;7:62.
8. Clark MD, Higgins R, Gumber A, et al. 'A better way to measure choices', discrete choice experiment/ conjoint analysis studies in Nephrology—a literature review. Eur Med J Nephrol. 2013;1:52–9.
9. Marshall D, Bridges JF, Hauber AB, et al. Conjoint analysis applications in health—how are studies being designed and reported? An update on current practice in the published literature between 2005 and 2008. Patient. 2010;3:249–56.
10. Anderson JL, Adams CD, Antman EM, et al. 2011 ACCF/AHA focused update incorporated into the ACC/AHA 2007 guidelines for the management of patients with unstable angina/non-ST-elevation myocardial infarction: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. Circulation. 2011;123:e426–579.
11. Reed Johnson F, Lancsar E, Marshall D, et al. Constructing experimental designs for discrete-choice experiments: report of the ISPOR Conjoint Analysis Experimental Design Good Research Practices Task Force. Value Health. 2013;16:3–13.
12. Coast J, Horrocks S. Developing attributes and levels for discrete choice experiments using qualitative methods. J Health Serv Res Policy. 2007;12:25–30.
13. Fraenkel L. Conjoint analysis at the individual patient level: issues to consider as we move from a research to a clinical tool. Patient. 2008;1:251–3.
14. Lancsar E, Louviere J. Conducting discrete choice experiments to inform healthcare decision making: a user's guide. Pharmacoeconomics. 2008;26:661–77.
15. Coast J, Al-Janabi H, Sutton EJ, et al. Using qualitative methods for attribute development for discrete choice experiments: issues and recommendations. Health Econ. 2012;21:730–41.
16. Bridges JF, Hauber AB, Marshall D, et al. Conjoint analysis applications in health—a checklist: a report of the ISPOR Good Research Practices for Conjoint Analysis Task Force. Value Health. 2011;14:403–13.
17. Flynn TN. Valuing citizen and patient preferences in health: recent developments in three types of best–worst scaling. Expert Rev Pharmacoecon Outcomes Res. 2010;10:259–67.
18. Lancsar E, Louviere J, Donaldson C, et al. Best worst discrete choice experiments in health: methods and an application. Soc Sci Med. 2013;76:74–82.
19. Bell RA, Paterniti DA, Azari R, et al. Encouraging patients with depressive symptoms to seek care: a mixed methods approach to message development. Patient Educ Couns. 2010;78:198–205.
20. De Achaval S, Fraenkel L, Volk RJ, et al. Impact of educational and patient decision aids on decisional conflict associated with total knee arthroplasty. Arthritis Care Res. 2012;64:229–37.
21. Fraenkel L, Chodkowski D, Lim J, et al. Patients' preferences for treatment of hepatitis C. Med Decis Making. 2010;30:45–57.
22. Fraenkel L. Feasibility of using modified adaptive conjoint analysis importance questions. Patient. 2010;3:209–15.

23. Gregorian RS Jr, Gasik A, Kwong WJ, et al. Importance of side effects in opioid treatment: a trade-off analysis with patients and physicians. J Pain. 2010;11:1095–108.
24. De Groot IB, Otten W, Smeets HJ, et al. Is the impact of hospital performance data greater in patients who have compared hospitals? BMC Health Serv Res. 2011;11:214.
25. De Groot IB, Otten W, Dijs-Elsinga J, et al. Choosing between hospitals: the influence of the experiences of other patients. Med Decis Making. 2012;32:764–78.
26. Meister H, Grugel L, Walger M, et al. Utility and importance of hearing-aid features assessed by hearing-aid acousticians. Trends Amplif. 2010;14:155–63.
27. Pieterse AH, Berkers F, Baas-Thijssen MCM, et al. Adaptive conjoint analysis as individual preference assessment tool: feasibility through the internet and reliability of preferences. Patient Educ Couns. 2010;78:224–33.
28. Pieterse AH, Stiggelbout AM, Marijnen CA. Methodologic evaluation of adaptive conjoint analysis to assess patient preferences: an application in oncology. Health Expect. 2010;13: 392–405.
29. Rochon D, Ebert JM, Fraenkel L, et al. Elderly patients' experiences using adaptive conjoint analysis software as a decision aid for osteoarthritis of the knee. Health Expect. 2012. [Epub]
30. Halme M, Linden K, Kaaria K. Patients' preferences for generic and branded over-the-counter medicines: an adaptive conjoint analysis approach. Patient. 2009;2:243–55.
31. Ahmed SF, Smith WA, Blamires C. Facilitating and understanding the family's choice of injection device for growth hormone therapy by using conjoint analysis. Arch Dis Child. 2008;93:110–4.
32. Beusterien KM, Dziekan K, Flood E, et al. Understanding patient preferences for HIV medications using adaptive conjoint analysis: feasibility assessment. Value Health. 2005;8:453–61.
33. Beusterien KM, Dziekan K, Schrader S, et al. Patient preferences among third agent HIV medications: a US and German perspective. AIDS Care. 2007;19:982–8.
34. Gan TJ, Lubarsky DA, Flood EM, et al. Patient preferences for acute pain treatment. Br J Anaesth. 2004;92:681–8.
35. Fraenkel L, Bodardus S, Wittnik DR. Understanding patient preferences for the treatment of lupus nephritis with adaptive conjoint analysis. Med Care. 2001;39:1203–16.
36. Fraenkel L, Bogardus ST Jr, Wittink DR. Risk-attitude and patient treatment preferences. Lupus. 2003;12:370–6.
37. Fraenkel L, Wittink DR, Concato J, et al. Informed choice and the widespread use of antiinflammatory drugs. Arthritis Rheum. 2004;51:210–4.
38. Fraenkel L, Bogardus ST, Concato J, et al. Patient preferences for treatment of rheumatoid arthritis. Ann Rheum Dis. 2004;63:1372–8.
39. Fraenkel L, Wittink DR, Concato J, et al. Are preferences for cyclooxygenase-2 inhibitors influenced by the certainty effect? J Rheumatol. 2004;31:591–3.
40. Fraenkel L, Gulanski B, Wittink D. Patient treatment preferences for osteoporosis. Arthritis Rheum. 2006;55:729–35.
41. Fraenkel L, Gulanski B, Wittink D. Patient willingness to take teriparatide. Patient Educ Couns. 2007;65:237–44.
42. Fraenkel L, Fried T. If you want patients with knee osteoarthritis (OA) to exercise: tell them about NSAIDS. Patient. 2008;1:21–6.
43. Pieterse AH, Stiggelbout AM, Baas-Thijssen MC, et al. Benefit from preoperative radiotherapy in rectal cancer treatment: disease-free patients' and oncologists' preferences. Br J Cancer. 2007;97:717–24.
44. Stiggelbout AM, de Vogel-Voogt E, Noordijk EM, et al. Individual quality of life: adaptive conjoint analysis as an alternative for direct weighting? Qual Life Res. 2008;17:641–9.
45. Boonen LH, Schut FT, Donkers B, et al. Which preferred providers are really preferred? Effectiveness of insurers' channeling incentives on pharmacy choice. Int J Health Care Finance Econ. 2009;9:347–66.
46. Damman OC, Hendriks M, Rademakers J, et al. Consumers' interpretation and use of comparative information on the quality of health care: the effect of presentation approaches. Health Expect. 2012;15:197–211.

47. Davison SN, Kromm SK, Currie GR. Patient and health professional preferences for organ allocation and procurement, end-of-life care and organization of care for patients with chronic kidney disease using a discrete choice experiment. Nephrol Dial Transplant. 2010;25:2334–41.

48. Eisingerich AB, Wheelock A, Gomez GB, et al. Attitudes and acceptance of oral and parenteral HIV preexposure prophylaxis among potential user groups: a multinational study. PLoS ONE. 2012;7:e28238.

49. Gerard K, Tinelli M, Latter S, et al. Valuing the extended role of prescribing pharmacist in general practice: results from a discrete choice experiment. Value Health. 2012;15:699–707.

50. Gidengil CT, Lieu TA, Payne K, et al. Parental and societal values for the risks and benefits of childhood combination vaccines. Vaccine. 2012;30:3445–52.

51. Goodall S, King M, Ewing J, et al. Preferences for support services among adolescents and young adults with cancer or a blood disorder: a discrete choice experiment. Health Policy. 2012;107:304–11.

52. Hancock-Howard RL, Ungar WJ, Marshall D, et al. Public preferences for counseling regarding antidepressant use during pregnancy: a discrete choice experiment. Birth Defects Res A Clin Mol Teratol. 2012;94:532–9.

53. Hill M, Fisher J, Chitty LS, et al. Women's and health professionals' preferences for prenatal tests for Down syndrome: a discrete choice experiment to contrast noninvasive prenatal diagnosis with current invasive tests. Genet Med. 2012;14:905–13.

54. Kimman ML, Dellaert BGC, Boersma LJ, et al. Follow-up after treatment for breast cancer: one strategy fits all? An investigation of patient preferences using a discrete choice experiment. Acta Oncol. 2010;49:328–37.

55. Kruk ME, Paczkowski MM, Tegegn A, et al. Women's preferences for obstetric care in rural Ethiopia: a population-based discrete choice experiment in a region with low rates of facility delivery. J Epidemiol Community Health. 2010;64:984–8.

56. Landfeldt E, Jablonowska B, Norlander E, et al. Patient preferences for characteristics differentiating ovarian stimulation treatments. Hum Reprod. 2012;27:760–9.

57. Mentzakis E, Ryan M, McNamee P. Using discrete choice experiments to value informal care tasks: exploring preference heterogeneity. Health Econ. 2011;20:930–44.

58. Miners A, Llewellyn C, Pollard A, et al. Assessing user preferences for sexually transmitted infection testing services: a discrete choice experiment. Sex Transm Infect. 2012;88:510–6.

59. Mohamed AF, Kilambi V, Luo MP, et al. Patient and parent preferences for immunoglobulin treatments: a conjoint analysis. J Med Econ. 2012;15:1183–91.

60. Naik-Panvelkar P, Armour C, Rose JM, et al. Patients' value of asthma services in Australian pharmacies: the way ahead for asthma care. J Asthma. 2012;49:310–6.

61. Naik-Panvelkar P, Armour C, Rose JM, et al. Patient preferences for community pharmacy asthma services: a discrete choice experiment. Pharmacoeconomics. 2012;30:961–76.

62. Pedersen LB, Kjaer T, Kragstrup J, et al. Do general practitioners know patients' preferences? An empirical study on the agency relationship at an aggregate level using a discrete choice experiment. Value Health. 2012;15:514–23.

63. Poulos C, Yang J, Patil SR, et al. Consumer preferences for household water treatment products in Andhra Pradesh, India. Soc Sci Med. 2012;75:738–46.

64. Van der Pol M, Shiell A, Au F, et al. Eliciting individual preferences for health care: a case study of perinatal care. Health Expect. 2010;13:4–12.

65. Waltzman JT, Scholz T, Evans GR. What patients look for when choosing a plastic surgeon: an assessment of patient preference by conjoint analysis. Ann Plast Surg. 2011;66:643–7.

66. Yeo ST, Edwards RT, Fargher EA, et al. Preferences of people with diabetes for diabetic retinopathy screening: a discrete choice experiment. Diabet Med. 2012;29:869–77.

67. Yi D, Ryan M, Campbell S, et al. Using discrete choice experiments to inform randomised controlled trials: an application to chronic low back pain management in primary care. Eur J Pain. 2011;15:531e1–10.

68. Bederman S, Mahomed N. In the eye of the beholder: preferences of patients, family physicians, and surgeons for lumbar spinal surgery. Spine. 2009;35(1):108–15.

69. Bridges JF, Mohamed AF, Finnern HW, et al. Patients' preferences for treatment outcomes for advanced non-small cell lung cancer: a conjoint analysis. Lung Cancer. 2012;77:224–31.

70. Chancellor J, Martin M, Liedgens H, et al. Stated preferences of physicians and chronic pain sufferers in the use of classic strong opioids. Value Health. 2012;15:106–17.

71. Clark MD, Gumber A, Leech D, et al. Prioritizing patients for renal transplantation? Analysis of patient preferences for kidney allocation according to ethnicity and gender. J Divers Health Soc Care. 2009;6:181–91.

72. Clark MD, Leech D, Gumber A, et al. Who should be prioritized for renal transplantation? Analysis of key stakeholder preferences using discrete choice experiments. BMC Nephrol. 2012;13:152.

73. Hauber AB, Mohamed AF, Reed Johnson F, et al. Treatment preferences and medication adherence of people with type 2 diabetes using oral glucose-lowering agents. Diabet Med. 2009;26(4):416–24.

74. Hauber AB, Mohamed AF, Reed Johnson F, et al. Estimating importance weights for the IWQOL-Lite using conjoint analysis. Qual Life Res. 2010;19:701–9.

75. Hauber AB, Gonzalez JM, Coombs J, et al. Patient preferences for reducing toxicities of treatments for gastrointestinal stromal tumor (GIST). Patient Prefer Adherence. 2011;5:307–14.

76. Reed Johnson F, Hauber AB, Ozdemir S. Using conjoint analysis to estimate healthy-year equivalents for acute conditions: an application to vasomotor symptoms. Value Health. 2009;12:146–52.

77. Potoglou D, Burge P, Flynn T, et al. Best–worst scaling vs. discrete choice experiments: an empirical comparison using social care data. Soc Sci Med. 2011;72:1717–27.

78. Ratcliffe J, Brazier J, Tsuchiya A, et al. Using DCE and ranking data to estimate cardinal values for health states for deriving a preference-based single index from the sexual quality of life questionnaire. Health Econ. 2009;18:1261–76.

79. Van Til JA, Stiggelbout AM, Ijzerman MJ. The effect of information on preferences stated in a choice-based conjoint analysis. Patient Educ Couns. 2009;74:264–71.

80. Wittink MN, Cary M, Tenhave T, et al. Towards patient-centered care for depression: conjoint methods to tailor treatment based on preferences. Patient. 2010;3:145–57.

81. Ahmed A, Fincham JE. Patients' view of retail clinics as a source of primary care: boon for nurse practitioners? J Am Acad Nurse Pract. 2011;23:193–9.

82. Albada A, Triemstra M. Patients' priorities for ambulatory hospital care centres. A survey and discrete choice experiment among elderly and chronically ill patients of a Dutch hospital. Health Expect. 2009;12:92–105.

83. Bansback N, Sizto S, Guh D, et al. The effect of direct-to-consumer genetic tests on anticipated affect and health-seeking behaviors: a pilot survey. Genet Test Mol Biomarkers. 2012;16:1165–71.

84. Bijlenga D, Bonsel GJ, Birnie E. Eliciting willingness to pay in obstetrics: comparing a direct and an indirect valuation method for complex health outcomes. Health Econ. 2011;20:1392–406.

85. Bogelund M, Visboll T, Faber J, et al. Patient preferences for diabetes management among people with type 2 diabetes in Denmark—a discrete choice experiment. Curr Med Res Opin. 2011;27:2175–83.

86. Bridges JF, Kinter ET, Schemding A, et al. Can patients diagnosed with schizophrenia complete choice-based conjoint analysis tasks? Patient. 2011;4:267–75.

87. Bridges JF, Lataille AT, Buttorff C, et al. Consumer preferences for hearing aid attributes: a comparison of rating and conjoint analysis methods. Trends Amplif. 2012;1640–8.

88. Bridges JF, Searle SC, Selck FW, et al. Designing family-centered male circumcision services: a conjoint analysis approach. Patient. 2012;5:101–11.

89. Brown DS, Finkelstein EA, Brown DR, et al. Estimating older adults' preferences for walking programs via conjoint analysis. Am J Prev Med. 2009; 36:201–7 e4.

90. Brown TM, Pashos CL, Joshi AV, et al. The perspective of patients with haemophilia with inhibitors and their care givers: preferences for treatment characteristics. Haemophilia. 2011;17:476–82.

91. Bunge E, de Bekker-Grob EW, van Biezen FC, et al. Patients' preferences for scoliosis brace treatment. Spine. 2009;35:57–63.

92. Burnett HF, Regier DA, Feldman BM, et al. Parents' preferences for drug treatments in juvenile idiopathic arthritis: a discrete choice experiment. Arthritis Care Res (Hoboken). 2012;64:1382–91.

93. Chan YM, Sahota DS, Leung TY, et al. Chinese women's preferences for prenatal diagnostic procedure and their willingness to trade between procedures. Prenat Diagn. 2009;29:1270–6.
94. Cheng J, Pullenayegum E, Marshall DA, et al. An empirical comparison of methods for analyzing correlated data from a discrete choice survey to elicit patient preference for colorectal cancer screening. BMC Med Res Methodol. 2012;12:15.
95. Damen TH, de Bekker-Grob EW, Mureau MA, et al. Patients' preferences for breast reconstruction: a discrete choice experiment. J Plast Reconstr Aesthet Surg. 2011;64:75–83.
96. Darba J, Restovic G, Kaskens L, et al. Patient preferences for osteoporosis in Spain: a discrete choice experiment. Osteoporos Int. 2011;22:1947–54.
97. De Bekker-Grob EW, Essink-Bot ML, Meerding WJ, et al. Preferences of GPs and patients for preventive osteoporosis drug treatment: a discrete-choice experiment. Pharmacoeconomics. 2009;27:211–9.
98. De Bekker-Grob EW, Hofman R, Donkers B, et al. Girls' preferences for HPV vaccination: a discrete choice experiment. Vaccine. 2010;28:6692–7.
99. Deverill M, Lancsar E, Snaith VB, et al. Antenatal care for first time mothers: a discrete choice experiment of women's views on alternative packages of care. Eur J Obstet Gynecol Reprod Biol. 2010;151:33–7.
100. Eberth B, Watson V, Ryan M, et al. Does one size fit all? Investigating heterogeneity in men's preferences for benign prostatic hyperplasia treatment using mixed logit analysis. Med Decis Making. 2009;29:707–15.
101. Essers BA, Dirksen CD, Prinsen MH, et al. Assessing the public's preference for surgical treatment of primary basal cell carcinoma: a discrete-choice experiment in the south of the Netherlands. Dermatol Surg. 2010;36:1950–5.
102. Essers BA, van Helvoort-Postulart D, Prins MH, et al. Does the inclusion of a cost attribute result in different preferences for the surgical treatment of primary basal cell carcinoma? A comparison of two discrete-choice experiments. Pharmacoeconomics. 2010;28:507–20.
103. Faggioli G, Scalone L, Mantovani LG, et al. Preferences of patients, their family caregivers and vascular surgeons in the choice of abdominal aortic aneurysms treatment options: the PREFER study. Eur J Vasc Endovasc Surg. 2011;42:26–34.
104. Glenngård AH, Hjelmgren J, Thomsen PH, et al. Patient preferences and willingness-to-pay for ADHD treatment with stimulants using discrete choice experiment (DCE) in Sweden, Denmark and Norway. Nord J Psychiatry. 2013;67:351–9.
105. Goto R, Takahashi Y. Nishimura S, et al. A cohort study to examine whether time and risk preference is related to smoking cessation success. Addiction. 2009;104:1018–24.
106. Goto R, Takahashi Y, Ida T. Changes in smokers' attitudes toward intended cessation attempts in Japan. Value Health. 2011;14:785–91.
107. Green C, Gerard K. Exploring the social value of health-care interventions: a stated preference discrete choice experiment. Health Econ. 2009;18:951–76.
108. Guimaraes C, Marra CA, Colley L, et al. A valuation of patients' willingness-to-pay for insulin delivery in diabetes. Int J Technol Assess Health Care. 2009;25:359–66.
109. Hodgkins P, Swinburn P, Solomon D, et al. Patient preferences for first-line oral treatment for mild-to-moderate ulcerative colitis: a discrete-choice experiment. Patient. 2012;5:33–44.
110. Hol L, de Bekker-Grob EW, van Dam L, et al. Preferences for colorectal cancer screening strategies: a discrete choice experiment. Br J Cancer. 2010;102:972–80.
111. Hong SH, Liu J, Wang J, et al. Conjoint analysis of patient preferences on Medicare medication therapy management. J Am Pharm Assoc (2003). 2011;51:378–87.
112. Howard K, Salkeld G. Does attribute framing in discrete choice experiments influence willingness to pay? Results from a discrete choice experiment in screening for colorectal cancer. Value Health. 2009;12:354–63.
113. Howard K, Salkeld G, Pignone M, et al. Preferences for CT colonography and colonoscopy as diagnostic tests for colorectal cancer: a discrete choice experiment. Value Health. 2011;14:1146–52.
114. Ijzerman MJ, van Til JA, Bridges JF. A comparison of analytic hierarchy process and conjoint analysis methods in assessing treatment alternatives for stroke rehabilitation. Patient. 2012;5:45–56.

115. Reed Johnson F, Ozdemir S, Phillips KA. Effects of simplifying choice tasks on estimates of taste heterogeneity in stated-choice surveys. Soc Sci Med. 2010;70:183–90.
116. Kauf TL, Mohamed AF, Hauber AB, et al. Patients' willingness to accept the risks and benefits of new treatments for chronic hepatitis C virus infection. Patient. 2012;5:265–78.
117. Kinter ET, Prior TJ, Carswell CI, et al. A comparison of two experimental design approaches in applying conjoint analysis in patient-centered outcomes research: a randomized trial. Patient. 2012;5:279–94.
118. Kiiskinen U, Suominen-Taipale AL, Cairns J. Think twice before you book? Modelling the choice of public vs private dentist in a choice experiment. Health Econ. 2010;19:670–82.
119. Koopmanschap MA, Stolk EA, Koolman X. Dear policy maker: have you made up your mind? A discrete choice experiment among policy makers and other health professionals. Int J Technol Assess Health Care. 2010;26:198–204.
120. Kruijshaar ME, Essink-Bot ML, Donkers B, et al. A labelled discrete choice experiment adds realism to the choices presented: preferences for surveillance tests for Barrett esophagus. BMC Med Res Methodol. 2009;9:31.
121. Laba TL, Brien JA, Jan S. Understanding rational non-adherence to medications. A discrete choice experiment in a community sample in Australia. BMC Fam Pract. 2012;13:61.
122. Lagarde M, Smith Paintain L. Evaluating health workers' potential resistance to new interventions: a role for discrete choice experiments. PLoS ONE. 2011;6:e23588.
123. Laver K, Ratcliffe J, George S, et al. Early rehabilitation management after stroke: what do stroke patients prefer? J Rehabil Med. 2011;43:354–8.
124. De Bekker-Grob EW, Rose JM, Bliemer MC. A closer look at decision and analyst error by including nonlinearities in discrete choice models: implications on willingness-to-pay estimates derived from discrete choice data in healthcare. Pharmacoeconomics. 2013;31:1169–83.
125. Lloyd A, Hodgkins P, Dewilde S, et al. Methylphenidate delivery mechanisms for the treatment of children with attention deficit hyperactivity disorder: heterogeneity in parent preferences. Int J Technol Assess Health Care. 2011;27:215–23.
126. Lloyd A, Nafees B, Barnett AH, et al. Willingness to pay for improvements in chronic long-acting insulin therapy in individuals with type 1 or type 2 diabetes mellitus. Clin Ther. 2011;33:1258–67.
127. Manjunath R, Yang JC, Ettinger AB. Patients' preferences for treatment outcomes of add-on antiepileptic drugs: a conjoint analysis. Epilepsy Behav. 2012;24:474–9.
128. Marti J. Assessing preferences for improved smoking cessation medications: a discrete choice experiment. Eur J Health Econ. 2012;13:533–48.
129. Mentzakis E, Stefanowska P, Hurley J. A discrete choice experiment investigating preferences for funding drugs used to treat orphan diseases: an exploratory study. Health Econ Policy Law. 2011;6:405–33.
130. Mohamed AF, Epstein JD, Li-McLeod JM. Patient and parent preferences for haemophilia A treatments. Haemophilia. 2011;17:209–14.
131. Morton RL, Snelling P, Webster AC, et al. Factors influencing patient choice of dialysis versus conservative care to treat end-stage kidney disease. CMAJ. 2012;184:E277–83.
132. Morton RL, Snelling P, Webster AC, et al. Dialysis modality preference of patients with CKD and family caregivers: a discrete-choice study. Am J Kidney Dis. 2012;60:102–11.
133. Muhlbacher AC, Nubling M. Analysis of physicians' perspectives versus patients' preferences: direct assessment and discrete choice experiments in the therapy of multiple myeloma. Eur J Health Econ. 2011;12:193–203.
134. Muhlbacher AC, Rudolph I, Lincke HJ, et al. Preferences for treatment of attention-deficit/hyperactivity disorder (ADHD): a discrete choice experiment. BMC Health Serv Res. 2009;9:149.
135. Musters AM, de Bekker-Grob EW, Mochtar MH, et al. Women's perspectives regarding subcutaneous injections, costs and live birth rates in IVF. Hum Reprod. 2011;26:2425–31.
136. Nayaradou M, Berchi C, Dejardin O, et al. Eliciting population preferences for mass colorectal cancer screening organization. Med Decis Making. 2010;30:224–33.
137. Nieboer AP, Koolman X, Stolk EA. Preferences for long-term care services: willingness to pay estimates derived from a discrete choice experiment. Soc Sci Med. 2010;70:1317–25.

138. Ozdemir S, Reed Johnson F, Hauber AB. Hypothetical bias, cheap talk, and stated willingness to pay for health care. J Health Econ. 2009;28:894–901.

139. Park MH, Jo C, Bae EY, et al. A comparison of preferences of targeted therapy for metastatic renal cell carcinoma between the patient group and health care professional group in South Korea. Value Health. 2012;15:933–9.

140. Pavlova M, Hendrix M, Nouwens E, et al. The choice of obstetric care by low-risk pregnant women in the Netherlands: implications for policy and management. Health Policy. 2009;93:27–34.

141. Pereira CC, Mulligan M, Bridges JF, et al. Determinants of influenza vaccine purchasing decision in the US: a conjoint analysis. Vaccine. 2011;29:1443–7.

142. Pignone MP, Brenner AT, Hawley S, et al. Conjoint analysis versus rating and ranking for values elicitation and clarification in colorectal cancer screening. J Gen Intern Med. 2012;27:45–50.

143. Regier DA, Diorio C, Ethier MC, et al. Discrete choice experiment to evaluate factors that influence preferences for antibiotic prophylaxis in pediatric oncology. PLoS ONE. 2012;7:e47470.

144. Ryan M, Watson V. Comparing welfare estimates from payment card contingent valuation and discrete choice experiments. Health Econ. 2009;18:389–401.

145. Scalone L, Mantovani LG, Borghetti F, et al. Patients', physicians', and pharmacists' preferences towards coagulation factor concentrates to treat haemophilia with inhibitors: results from the COHIBA Study. Haemophilia. 2009;15:473–86.

146. Scalone L, Watson V, Ryan M, et al. Evaluation of patients' preferences for genital herpes treatment. Sex Transm Dis. 2011;38:802–7.

147. Schaarschmidt ML, Schmieder A, Umar N, et al. Patient preferences for psoriasis treatments: process characteristics can outweigh outcome attributes. Arch Dermatol. 2011;147:1285–94.

148. Schwappach DL, Mülders V, Simic D, et al. Is less more? Patients' preferences for drug information leaflets. Pharmacoepidemiol Drug Saf. 2011;20:987–95.

149. Scotland GS, McNamee P, Cheyene H, et al. Women's preferences for aspects of labor management: results from a discrete choice experiment. Birth. 2011;38:36–46.

150. Slothuus Skjoldborg U, Lauridsen J, Junker P. Reliability of the discrete choice experiment at the input and output level in patients with rheumatoid arthritis. Value Health. 2009;12:153–8.

151. Sung L, Alibhai SM, Ethier M, et al. Discrete choice experiment produced estimates of acceptable risks of therapeutic options in cancer patients with febrile neutropenia. J Clin Epidemiol. 2012;65:627–34.

152. Sweeting KR, Whitty JA, Scuffham PA, et al. Patient preferences for treatment of Achilles tendon pain: results from a discrete-choice experiment. Patient. 2011;4:45–54.

153. Swinburn P, Lloyd A, Ali S, et al. Preferences for antimuscarinic therapy for overactive bladder. BJU Int. 2011;108:868–73.

154. Thrumurthy SG, Morris JJ, Mughal MM, et al. Discrete-choice preference comparison between patients and doctors for the surgical management of oesophagogastric cancer. Br J Surg. 2011;98:1124-31 (discussion 1132).

155. Tinelli M, Ryan M, Bond C. Patients' preferences for an increased pharmacist role in the management of drug therapy. Int J Pharm Pract. 2009;17:275–82.

156. Tinelli M, Ozolins M, Bath-Hextall F, et al. What determines patient preferences for treating low risk basal cell carcinoma when comparing surgery vs imiquimod? A discrete choice experiment survey from the SINS trial. BMC Dermatol. 2012;12:19.

157. Van Dam L, Hol L, de Bekker-Grob EW, et al. What determines individuals' preferences for colorectal cancer screening programmes? A discrete choice experiment. Eur J Cancer. 2010;46:150–9.

158. Van der Pol M, McKenzie L. Costs and benefits of tele-endoscopy clinics in a remote location. J Telemed Telecare. 2010;16:89–94.

159. Van Empel IW, Dancet EA, Koolman XH, et al. Physicians underestimate the importance of patient-centredness to patients: a discrete choice experiment in fertility care. Hum Reprod. 2011;26:584–93.

160. Van Houtven G, Reed Johnson F, Kilambi V, et al. Eliciting benefit-risk preferences and probability-weighted utility using choice-format conjoint analysis. Med Decis Making. 2011;31:469–80.

161. Witt J, Scott A, Osborne RH. Designing choice experiments with many attributes. An application to setting priorities for orthopaedic waiting lists. Health Econ. 2009;18:681–96.

162. Wong MK, Mohamed AF, Hauber AB, et al. Patients rank toxicity against progression free survival in second-line treatment of advanced renal cell carcinoma. J Med Econ. 2012;15:1139–48.

163. Bansback N, Brazier J, Tsuchiya A, et al. Using a discrete choice experiment to estimate health state utility values. J Health Econ. 2012;31:306–18.

164. Stolk EA, Oppe M, Scalone L, et al. Discrete choice modeling for the quantification of health states: the case of the EQ-5D. Value Health. 2010;13:1005–13.

165. Lancsar E, Wildman J, Donaldson C, et al. Deriving distributional weights for QALYs through discrete choice experiments. J Health Econ. 2011;30: 466–78.

166. Van der Wulp I, van den Hout WB, de Vries M, et al. Societal preferences for standard health insurance coverage in the Netherlands: a cross-sectional study. BMJ Open. 2012;2:e001021.

167. Blaauw D, Erasmus E, Pagaiya N, et al. Policy interventions that attract nurses to rural areas: a multicountry discrete choice experiment. Bull World Health Organ. 2010;88:350–6.

168. Grindrod KA, Marra CA, Colley L, et al. Pharmacists' preferences for providing patient-centered services: a discrete choice experiment to guide health policy. Ann Pharmacother. 2010;44:1554–64.

169. Gunther OH, Kürstein B, Riedel-Heller SG, et al. The role of monetary and nonmonetary incentives on the choice of practice establishment: a stated preference study of young physicians in Germany. Health Serv Res. 2010;45:212–29.

170. Huicho L, Miranda JJ, Diez-Canseco F, et al. Job preferences of nurses and midwives for taking up a rural job in Peru: a discrete choice experiment. PLoS ONE. 2012;7:e50315.

171. Kolstad JR. How to make rural jobs more attractive to health workers. Findings from a discrete choice experiment in Tanzania. Health Econ. 2011;20:196–211.

172. Miranda JJ, Diez-Canseco F, Lema C, et al. Stated preferences of doctors for choosing a job in rural areas of Peru: a discrete choice experiment. PLoS ONE. 2012;7:e50567.

173. Rockers PC, Jaskiewicz W, Wurts L, et al. Preferences for working in rural clinics among trainee health professionals in Uganda: a discrete choice experiment. BMC Health Serv Res. 2012;12:212.

174. Sivey P, Scott A, Witt J, et al. Junior doctors' preferences for specialty choice. J Health Econ. 2012;31:813–23.

175. Carlsen B, Hole AR, Kolstad JR, et al. When you can't have the cake and eat it too: a study of medical doctors' priorities in complex choice situations. Soc Sci Med. 2012;75 1964–73.

176. Defechereux T, Paolucci F, Mirelman A, et al. Health care priority setting in Norway a multicriteria decision analysis. BMC Health Serv Res. 2012;12:39.

177. Diederich A, Swait J, Wirsik N. Citizen participation in patient prioritization policy decisions: an empirical and experimental study on patients' characteristics. PLoS ONE. 2012;7:e36824.

178. Kjaer T, Bech M, Kronborg C, et al. Public preferences for establishing nephrology facilities in Greenland: estimating willingness-to-pay using a discrete choice experiment. Eur J Health Econ. 2013;14:739–48.

179. Lim MK, Bae EY, Choi S, et al. Eliciting public preference for health-care resource allocation in South Korea. Value Health. 2012;15:S91–4.

180. Marsh K, Dolan P, Kempster J, et al. Prioritizing investments in public health: a multi-criteria decision analysis. J Public Health. 2013;35:460–6.

181. Mirelman A, Mentzakis E, Kinter E, et al. Decision-making criteria among national policymakers in five countries: a discrete choice experiment eliciting relative preferences for equity and efficiency. Value Health. 2012;15:534–9.

182. Ng V, Sargeant JM. A quantitative and novel approach to the prioritization of zoonotic diseases in North America: a public perspective. PLoS One. 2012;7:e48519.

183. Philips H, Mahr D, Remmen R, et al. Predicting the place of out-of-hours care—a market simulation based on discrete choice analysis. Health Policy. 2012;106:284–90.

184. Promberger M, Dolan P, Marteau TM. ''Pay them if it works'': discrete choice experiments on the acceptability of financial incentives to change health related behaviour. Soc Sci Med. 2012;75:2509–14.

185. Rennie L, Porteous T, Ryan M. Preferences for managing symptoms of differing severity: a discrete choice experiment. Value Health. 2012;15:1069–76.

186. Scuffham PA, Whitty JA, Taylor M, et al. Health system choice: a pilot discrete-choice experiment eliciting the preferences of British and Australian citizens. Appl Health Econ Health Policy. 2010;8: 89–97.

187. Watson V, Carnon A, Ryan M, et al. Involving the public in priority setting: a case study using discrete choice experiments. J Public Health (Oxf). 2012;34:253–60.

188. Watson V, Sussex J, Ryan M, et al. Managing poorly performing clinicians: health care providers' willingness to pay for independent help. Health Policy. 2012;104:260–71.

189. Youngkong S, Baltussen R, Tantivess S, et al. Criteria for priority setting of HIV/AIDS interventions in Thailand: a discrete choice experiment. BMC Health Serv Res. 2010;10:197.

190. Arden NK, Hauber AB, Mohamed AF, et al. How do physicians weigh benefits and risks associated with treatments in patients with osteoarthritis in the United Kingdom? J Rheumatol. 2012;39:1056–63.

191. Benjamin L, Cotté FE, Philippe C, et al. Physicians' preferences for prescribing oral and intravenous anticancer drugs: a discrete choice experiment. Eur J Cancer. 2012;48:912–20.

192. Bhatt M, Currie GR, Auld MC, et al. Current practice and tolerance for risk in performing procedural sedation and analgesia on children who have not met fasting guidelines: a Canadian survey using a stated preference discrete choice experiment. Acad Emerg Med. 2010;17:1207–15.

193. Jackman J, Uy M, Hsieh Y, et al. Minding the gap: an approach to determine critical drivers in the development of point of care diagnostics. Point Care. 2012;11:130–9.

194. Mohamed AF, Reed Johnson F, Hauber AB, et al. Physicians' stated trade-off preferences for chronic hepatitis B treatment outcomes in Germany, France, Spain, Turkey, and Italy. Eur J Gastroenterol Hepatol. 2012;24:419–26.

195. Nathan H, Bridges JF, Cosgrove DP, et al. Treating patients with colon cancer liver metastasis: a nationwide analysis of therapeutic decision making. Ann Surg Oncol. 2012;19:3668–76.

196. Torbica A, Fattore G. Understanding the impact of economic evidence on clinical decision making: a discrete choice experiment in cardiology. Soc Sci Med. 2010;70:1536–43.

197. Tsung-Tai C, Heng-Chaing C, Lao-Nga M. Using discrete choice experiments to elicit doctors' preferences for report card design of diabetes care in Taiwan—a pilot study. J Eval Clin Pract. 2010;16:14–20.

198. Van Helvoort-Postulart D, Dellaert BG, van der Weijden T, et al. Discrete choice experiments for complex health-care decisions: does hierarchical information integration offer a solution? Health Econ. 2009;18:903–20.

199. Van Helvoort-Postulart D, van der Weijden T, Dellaert BG, et al. Investigating the complementary value of discrete choice experiments for the evaluation of barriers and facilitators in implementation research: a questionnaire survey. Implement Sci. 2009;4:10.

200. Wyatt JC, Batley RP, Keen J. GP preferences for information systems: conjoint analysis of speed, reliability, access and users. J Eval Clin Pract. 2010;16:911–5.

201. Pedersen LB, Kjaer T, Kragstrup J, et al. General practitioners' preferences for the organisation of primary care: a discrete choice experiment. Health Policy. 2012;106:246–56.

202. Al Hamarneh YN, Agus A, Campbell D, et al. Public perceptions of coronary events risk factors: a discrete choice experiment. BMJ Open. 2012;2;e001560.

203. Bech M, Kjaer T, Lauridsen J. Does the number of choice sets matter? Results from a web survey applying a discrete choice experiment. Health Econ. 2011;20(3):273–86.

204. Cunningham CE, Henderson J, Niccols A, et al. Preferences for evidence-based practice dissemination in addiction agencies serving women: a discrete-choice conjoint experiment. Addiction. 2012;107:1512–24.

205. Fegert JM, Slawik L, Wermelskirchen D, et al. Assessment of parents' preferences for the treatment of school-age children with ADHD: a discrete choice experiment. Expert Rev Pharmacoecon Outcomes Res. 2011;11:245–52.

206. Oteng B, Marra F, Lynd LD, et al. Evaluating societal preferences for human papillomavirus vaccine and cervical smear test screening programme. Sex Transm Infect. 2011;87:52–7.

207. Regier DA, Friedman JM, Makela N, et al. Valuing the benefit of diagnostic testing for genetic causes of idiopathic developmental disability: willingness to pay from families of affected children. Clin Genet. 2009;75:514–21.
208. Robyn PJ, Bärnighausen T, Souares A, et al. Health worker preferences for community-based health insurance payment mechanisms: a discrete choice experiment. BMC Health Serv Res. 2012;12:159.
209. Scasny M, Alberini A. Valuation of mortality risk attributable to climate change: investigating the effect of survey administration modes on a VSL. Int J Environ Res Public Health. 2012;9:4760–81.
210. Schellings R, Essers BAB, Kessels AG, et al. The development of quality indicators in mental healthcare: a discrete choice experiment. BMC Psychiatry. 2012;12:103.
211. Tong BC, Huber JC, Ascheim DD, et al. Weighting composite endpoints in clinical trials: essential evidence for the heart team. Ann Thorac Surg. 2012;94:1908–13.
212. Tonin S, Alberini A, Turvani M. The value of reducing cancer risks at contaminated sites: are more knowledgeable people willing to pay more? Risk Anal. 2012;32:1157–82.
213. Vroomen J, Zweifel P. Preferences for health insurance and health status: does it matter whether you are Dutch or German? Eur J Health Econ. 2011;12:87–95.
214. Whitty JA, Scuffham PA, Rundle-Thiele SR. Public and decision maker stated preferences for pharmaceutical subsidy decisions: a pilot study. Appl Health Econ Health Policy. 2011;9:73–9.
215. Idkowiak J, Lavery GG, Dhir V, et al. Premature adrenarche: novel lessons from early onset androgen excess. Eur J Endocrinol. 2011;165:189–207.
216. Hole AR. Modelling heterogeneity in patients' preferences for the attributes of a general practitioner appointment. J Health Econ. 2008;27:1078–94.
217. Mark TL, Swait J. Using stated preference modeling to forecast the effect of medication attributes on prescriptions of alcoholism medications. Value Health. 2003;6:474–82.
218. Lancsar E, Louviere J. Deleting 'irrational' responses from discrete choice experiments: a case of investigating or imposing preferences? Health Econ. 2006;15:797–811.
219. Ryan M. Using conjoint analysis to take account of patient preferences and go beyond health outcomes: an application to in vitro fertilisation. Soc Sci Med. 1999;48:535–46.
220. Ryan M, Netten A, Skåtun D, et al. Using discrete choice experiments to estimate a preference-based measure of outcome—an application to social care for older people. J Health Econ. 2006;25:927–44.
221. Burr JM, Kilonzo M, Vale L, et al. Developing a preference-based Glaucoma Utility Index using a discrete choice experiment. Optom Vis Sci. 2007;84:797–808.
222. Slothuus Skjoldborg U, Gyrd-Hansen D. Conjoint analysis. The cost variable: an Achilles' heel? Health Econ. 2003;12:479–91.
223. Gyrd-Hansen D, Slothuus Skjoldborg U. The price proxy in discrete choice experiments: Issues of relevance for future research. In: Ryan M, Gerard K, Amaya-Amaya M, editors. Using discrete choice experiments to value health and health care; 2008. p. 175–193.
224. Louviere JJ, Lancsar E. Choice experiments in health: the good, the bad, the ugly and toward a brighter future. Health Econ Policy Law. 2009;4:527–46.
225. Mark TL, Swait J. Using stated preference and revealed preference modeling to evaluate prescribing decisions. Health Econ. 2004;13:563–73.
226. Reed Johnson F, Mohamed AF, Ozdemir S, et al. How does cost matter in health-care discrete-choice experiments? Health Econ. 2011;20:323–30.
227. Ratcliffe J, Buxton M, McGarry T, et al. Patients' preferences for characteristics associated with treatments for osteoarthritis. Rheumatology (Oxford). 2004;43:337–45.
228. Deal K. Segmenting patients and physicians using preferences from discrete choice experiments. Patient. 2014;7:5–21.
229. Miguel FS, Ryan M, Amaya-Amaya M. 'Irrational' stated preferences: a quantitative and qualitative investigation. Health Econ. 2005;14:307–22.

ADDITIONAL FILE 1: THE USE OF ADAPTIVE CONJOINT ANALYSIS (ACA) / AND ADAPTIVE CHOICE BASED CONJOINT (ACBC) STUDIES.

Some trends that can be discerned in relation to the Adaptive Conjoint Analysis (ACA) / and Adaptive Choice Based Conjoint (ACBC) studies are apparent from the data in Table A1.

Table A1 | Background information on ACA / ACBC studies.

Item	Category	Period: 2001-2008 (N=14)	Period: 2009-2012 (N=12)
Country of origin	UK	1 (7.1%)	0 (0%)
	USA	11 (78.6%)	6 (50%)
	Australia	0 (0%)	0 (0%)
	Canada	0 (0%)	0 (0%)
	Denmark	0 (0%)	0 (0%)
	Netherlands	2 (14.3%)	4 (33.3%)
	Germany	1 (7.1%)	1 (8.3%)
	Other	0 (0%)	2 (16.7%)
Number of attributes	2 – 3	0 (0%)	0 (0 %)
	4 – 5	6 (42.9%)	4 (33.3%)
	6	1 (7.1%)	3 (25%)
	7 – 9	2 (14.3%)	2 (16.7%)
	10	0 (0%)	1 (8.33%)
	>10	4 (28.6%)	2 (16.7%)
	Not clearly reported	1 (7.1%)	0 (0%)
Attributes covered	Monetary measure	2 (14.3%)	4 (25%)
	Time	8 (57.1%)	6 (50%)
	Risk	11 (78.6%)	10 (83.3%)
	Health Status domain	14 (100%)	12 (100%)
	Health care	12 (85.7%)	8 (66.7%)
	Other	1 (7.1%)	3 (25%)
Number of choices per respondent	8 or less	3 (21.4%)	2 (16.7%)
	9 – 16 choices	2 (14.3%)	4 (33.2%)
	More than 16 choices	3 (21.4%)	6 (50%)
	Not clearly reported	6 (42.9%)	0 (0%)
Administration of survey	Self-completed questionnaire	0 (0%)	1 (8.3%)
	Interviewer administered	8 (57.1%)	0 (0%)
	Computerized review	6 (42.9%)	10 (83.3%)
	Not reported	0 (0%)	1 (8.3%)

Notes: Some of the categories in tables do not add up to 100%, because they are not necessarily mutually exclusive, and studies may include multiple categories.

Searches using PubMed using the search terms 'Adaptive Conjoint Analysis', 'Adaptive Choice Based Conjoint', and 'menu-based conjoint' uncovered no healthcare related studies of this type for 1990-2000. For 2001-2008 however we obtained 14 analyses in 14 papers (equivalent to 1.75 per year). Moreover, for 2009-2012 we obtained and reviewed 12 analyses in 12 papers (equivalent to 3 per year). Many of these analyses have emanated from the USA (78.6% in 2001-2008, and 50% in 2009-2012), and the Netherlands (14.3% in 2001-2008, and 33.3% in 2009-2012). All of the analyses involved 4 or more attributes (where the number of attributes was specified). Moreover, some of the studies involved in excess of 10 attributes (28.6% in 2001-2008, and 16.7% in 2009-2012). The ACA / ACBC approach has the advantage that it can easily cater for designs with large numbers of attributes. Analyses often involved attributes which encompassed a wide range of domains (see Table A1). Like the DCE sample of analyses reviewed in the main paper, these analyses seemed to vary considerably in relation to the number of choices each respondent has faced. The analyses rarely seemed to involve completion of 'Self-completed questionnaires.' This is because they usually required the completion of computer-generated questionnaires. So they were usually either interviewer administered, with the interviewer inputting information into an ACA computer software package (occurred in 57.1% of analyses in 2001-2008; and 0% in 2009-2012) or administered by computerized review (occurred in 42.9% of analyses in 2001-2008, and 83.3% in 2009-2012).

The data in Table A2, indicates that when the design type is indicated (100% of cases in 2001-2008, and 83.3% of cases in 2009-2012), the use of a fractional factorial rather than full factorial design is used. There seems to be a major problem in terms of the analyses not indicating whether a main effects design or a design involving main effects plus interaction effects is used (100% of cases in 2001-2008, and 100% of cases in 2009-2012). The design source was usually said to be Sawtooth software (100% in 2001-2008, and 91.7% in 2009-2012), although one analysis in 2009-2012 (8.3%) just said it was an unnamed ACA software package.

Table A2 | Background information on ACA / ACBC studies.

Item	Category	Period: 2001-2008 (N=14)	Period: 2009-2012 (N=12)
Design type	Full factorial	0 (0%)	0 (0%)
	Fractional factorial	14 (100%)	10 (83.3%)
	Not clearly reported	0 (0%)	2 (16.7%)
Design plan	Main effects only	0 (0%)	0 (0%)
	Main effects, 2 or more way interactions	0 (0%)	0 (0%)
	Not applicable	0 (0%)	0 (0%)
	Not clearly reported	14 (100%)	12 (100%)
Design Source	Software package	14 (100%)	12 (100%)
	SPEED	0 (0%)	0 (0%)
	SPSS	0 (0%)	0 (0%)
	SAS	0 (0%)	0 (0%)
	SAWTOOTH	14 (100%)	11 (91.7%)
	Other	0 (0%)	0 (0%)
	No further details	0 (0%)	1 (8.3%)
	Catalogue	0 (0%)	0 (0%)
	Website	0 (0%)	0 (0%)
	Expert	0 (0%)	0 (0%)
	Not clearly reported	0 (0%)	1 (8.3%)
Method to create choice sets	Orthogonal arrays: Single profiles (i.e. binary choices)	0 (0%)	0 (0%)
	Orthogonal arrays: Random pairing	0 (0%)	0 (0%)
	Orthogonal arrays: Pairing with constant comparator	0 (0%)	0 (0%)
	Orthogonal arrays: Foldover – random pairing	0 (0%)	0 (0%)
	Orthogonal arrays: Foldover	0 (0%)	0(0%)
	D-efficiency	0 (0%)	0 (0%)
	Other (pragmatically chosen)	0 (0%)	0 (0%)
	Not clearly reported	0 (0%)	1 (8.3%)
	Other	14 (100%)	11 (91.7%)

Table A3 provides information on the range of estimation procedures used. These did not tend to fall into the main categories used for the data analysis for DCEs because typically other data analysis methods were used[8] (64.3% of analyses in 2001-2008, and 91.7% of analyses in 2009-2012), or there was a failure to report which methods were used (35.7% in 2001-2008 and 8.3% in 2009-2012) with respect to a number of analyses.

[8] The range of 'other' analyses used (n=9) which were indicated during 2001-2008 included 5 studies using ordinary least squares; 1 study using multiple logistic regression; 2 studies using Sawtooth Software generated utility values; and 1 study using Sawtooth Software to establish the percentage of patients preferring a treatment option. Moreover, the range of other analyses (n=11) indicated during 2009-2012 included 4 studies using OLS estimation; 1 study using a least squares updating algorithm; 3 studies using Hierarchical Bayesian estimation; 1 study reporting the use of Sawtooth Software with calculation of adjusted odds ratios; 1 study referring to use of Sawtooth ACA Software, and Hierarchical Bayesian and SMRT market simulation models; and 1 study which refers to use of ACA Software but did not define the package used.

Table A3 | Estimation procedures.

Estimation method	**Period: 2001-2008 (N=14)**	**Period: 2009-2012 (N=12)**
Probit	0 (0%)	0 (0%)
Random Effects Probit	0 (0%)	0 (0%)
Logit	0 (0%)	0 (0%)
Random Effects Logit	0 (0%)	0 (0%)
MNL	0 (0%)	0 (0%)
Nested Logit	0 (0%)	0 (0%)
Mixed Logit / Random Parameter Logit	0 (0%)	0 (0%)
Latent class (LCM)	0 (0%)	0 (0%)
Other	14 (100%)	11(93.7%)
Not clearly reported	0 (0%)	1 (8.3%)
Further details about Mixed Logit [MXL] / Random Parameter Logit Analyses [RPL]	**Period: 2001-2008 (N=14)**	**Period: 2009-2012 (N=12)**
Number described as RPL only	0 (0%)	0 (0%)
Number described as MXL only	0 (0%)	0 (0%)
Number described as both RPL and MXL analyses	0 (0%)	0 (0%)
Number of analyses uncovering evidence of preference heterogeneity	0 (0%)	0 (0%)
Number of papers indicating the number of replications (e.g. the number of Halton draws)	0 (0%)	0 (0%)
Number of analyses saying they used a normal distribution for random parameters	0 (0%)	0 (0%)
Number of analyses saying they used a log distribution for random parameters	0 (0%)	0 (0%)
Number of analyses saying they assumed another type of distribution for random parameters	0 (0%)	0 (0%)
Number of analyses failing to provide information about distributional assumptions of random parameters	0 (0%)	0 (0%)
Estimation procedure 'other'	9 (64.3%)	11 (91.7%)
Estimation procedure not clearly reported	5 (35.7%)	1 (8.3%)

Table A4 contains information on the use of validity checks. These did not seem to be deployed very often for these types of analysis. For the period 2001-2008 we found little evidence that these had been deployed. However, one analysis deployed a test for non-satiation (7.1%) and one analysis [1] undertook an external validity check (7.1%). For the period 2009-2012, 25% of analyses used validity checks to check for non-satiation. Moreover, 25% of analyses used qualitative analysis to inform attribute selection, and 8.3% of analyses used qualitative analysis to inform level selection.

Table A4 | Validity.

Item	Category	Period: 2001-2008 (N=14)	Period: 2009-2012 (N=12)
Validity tests	External	1 (7.1%)	0 (0%)
	Internal: Theoretical	0 (0%)	0 (0%)
	Internal: Non-satiation	1 (7.1%)	3 (25%)
	Internal: Transitivity	0 (0%)	0 (0%)
	Internal: Sen's expansion and contraction	0 (0%)	0 (0%)
	Internal: Compensatory decision making	0 (0%)	0 (0%)
Use of qualitative methods to enhance DCE process & results	Increasing face validity: Attribute selection	0 (%)	3 (25%)
	Increasing face validity: Level selection	0 (0%)	1 (8.3%)
	Increasing face validity: Pre-testing questionnaire	0 (0%)	0 (0%)
	Increasing face validity: Strengthen understanding responses – Debriefing choices	0 (0%)	0 (0%)

Finally, Table A5 provides information on the output of the ACA / ACBC analyses. During 2001-2008, 21.4% of analyses related to valuing health outcomes (Category B type analyses), and 78.6% of analyses related to valuing trade-offs between health outcomes and experience factors (Category C analyses). Moreover, a range of outcome measures were used (for details refer to Table A5 and the footnotes to the Table).

During 2009-2012, 25% of analyses related to valuing health outcomes (Category B type analyses), and 58.3% of analyses related to valuing trade-offs between health outcomes and experience factors (Category C analyses). One analysis (8.3%) related to valuing health professionals' preferences (Category G analysis) and another (8.3%) related to the category 'Other' (Category H analysis). Moreover, a range of outcome measures was used (for details refer to table A5 and the footnotes to the Table).

Table A5 | Output of ACA / ACBC analyses.

Main study objective	Number	Per WTP unit	Per time unit	Per risk unit	Monetary welfare measure	Utility score	Odds ratio	Probability score	Other
Period: 2001-2008 (N=14)									
(A) Valuing experience factors	0 (0%)	0	0	0	0	0	0	0	0
(B) Valuing health outcomes	3 (21.4%)	0	0	0	0	0	0	0	3[1]
(C) Trade-offs health outcomes & experience factors	11 (78.6%)	0	0	0	0	3	0	0	8[2]
(D) Utility weights within QALY framework	0 (0%)	0	0	0	0	0	0	0	0
(E) Job-choices	0 (0%)	0	0	0	0	0	0	0	0
(F) Developing priority setting frameworks	0 (0%)	0	0	0	0	0	0	0	0
(G) Health professional's preferences	0 (0%)	0	0	0	0	0	0	0	0
(H) Other	0 (0%)	0	0	0	0	0	0	0	0
Total	14 (100%)	0	0	0	0	3	0	0	11
Period: 2009-2012 (N=12)									
(A) Valuing experience factors	0 (0%)	0	0	0	0	0	0	0	0
(B) Valuing health outcomes	3 (25%)	0	0	0	0	0	0	0	3[3]
(C) Trade-offs health outcomes & experience factors	7 (58.3%)	2	0	0	0	0	0	1	4[4]
(D) Utility weights within QALY framework	0 (0%)	0	0	0	0	0	0	0	0
(E) Job-choices	0 (0%)	0	0	0	0	0	0	0	0
(F) Developing priority setting frameworks	0 (0%)	0	0	0	0	0	0	0	0
(G) Health professional's preferences	1 (8.3%)	0	0	0	0	0	0	0	1[5]
(H) Other	1 (8.3%)	0	0	0	0	0	0	0	1[6]
Total	12 (100%)	2	0	0	0	0	0	1	7

Notes: [1]'Other' included 1 analysis relating to relative importance of attributes; 1 analysis relating to 'predicted preferences for NSAIDs vs. Cox inhibitors'; and 1 analysis relating to computation of individual Quality of Life (iQoL) scores. [2]'Other' included 3 analyses using a measure involving 'relative importance of attributes'; 1 relating to 'Other medication treatment'; 1 relating to 'Other forms of medication'; 1 relating to 'Other respondents treatment choices'; 1 relating to 'Other preferences for taking Teriparatide'; and 1 relating to 'Other preferences for different interventions'. [3]'Other' involved 3 studies each using a relative importance score' for attributes endpoint. [4]'Other' involved 2 studies involving a 'relative importance score' for attributes; 1 study relating to differences in 'decisional conflict'; and 1 study involving relative percentage importance of attributes. [5]'Other' involves an analysis using a 'relative importance score' endpoint. [6]'Other' involves an analysis using a 'relative importance score' endpoint.

REFERENCE

1. Ahmed SF, Smith WA, Blamires C. Facilitating and understanding the family's choice of injection device for growth hormone therapy by using conjoint analysis. Arch Dis Child. 2008;93:110–4.

PART II

The present - state of the art applications

3

Future pandemics and vaccination: public opinion and attitudes across three European countries

Domino Determann
Esther W. de Bekker-Grob
Jeff French
Helene A. Voeten
Jan Hendrik Richardus
Enny Das
Ida J. Korfage

Vaccine. 2016;34(6):803-8
doi: 10.1016/j.vaccine.015.12.035

ABSTRACT

Background: Understanding public opinion and attitudes regarding vaccination is crucial for successful outbreak management and effective communication at the European level.

Methods: We explored national differences by conducting focus group discussions in the Netherlands, Poland and Sweden. Discussions were structured using concepts from behavioural models.

Results: Thematic analysis revealed that participants would base their vaccination decision on trade- offs between perceived benefits and barriers of the vaccine also taking into account the seriousness of the new outbreak. Except for those having chronic diseases, participants expected a low infection risk, resulting in a low willingness to get vaccinated. Information about the health status of cases was considered important since this might change perceived susceptibility. Participants displayed concerns about vaccine safety due to the limited available time to produce and test vaccines in the acute situation of a new pandemic. Swedish participants mentioned their tendency of doing the right thing and following the rules, as well as to get vaccinated because of solidarity with other citizens and social influences. This appeared much less prominent for the Dutch and Polish participants. However, Swedish participants indicated that their negative experiences during the Influenza A/H1N1 2009 pandemic decreases their acceptance of future vaccinations. Polish participants lacked trust in their national (public) health system and government, and were therefore sceptical about the availability and quality of vaccines in Poland.

Conclusions: Although participants overall expressed similar considerations, important differences between countries stand out, such as previous vaccination experiences, the degree of adherence to social norms, and the degree of trust in health authorities.

INTRODUCTION

Outbreaks of communicable diseases will cross borders, with Influenza A/H1N1 [1] and Ebola [2] being recent examples, and increased international travel and migration will facilitate their speed and spread [3]. Cross-border collaboration in the management of future outbreaks within Europe is therefore necessary. Since public health professionals and authorities will be focused on controlling the spread and impact of the new disease during such an outbreak [4], it is essential to timely update and improve existing European pandemic preparedness plans, preferably before outbreaks begin [5].

The success of mitigating a new outbreak is largely dependent on the willingness of the public to comply with recommended preventive measures. Understanding the public opinion and attitudes regarding preventive measures is thus crucial for successful outbreak management and effective communication. Reasons to accept or decline preventive measures in pandemic situations have been described [4, 6–12], but very little is known about potential differences herein across Europe.

We therefore conducted focus group discussions in three countries across Europe to explore (1) the public opinion and attitude regarding future pandemics and vaccination and (2) potential differences in opinions and attitudes between participants in the Netherlands, Sweden, and Poland.

METHODS

We opted for focus group discussions (FGDs) [13, 14] to explore public opinion and attitudes. FGDs were chosen because these enable unforeseen topics to arise and to be explored in depth [15]. We developed a theory based semi-structured question route based on the Health Belief Model and two elements from other behavioural models (additional files 1 and 2). The question route was pilot tested, evaluated and improved where necessary. The Medical Ethics Committee of the Erasmus University Medical Center Rotterdam approved the study protocol (MEC-2012-263). Independent research agencies recruited 6–9 participants per FGD and used purposive sampling methods to ensure a diverse sample regarding age, sex, and educational level. Participants received a financial incentive for their contribution, adapted to the national norm.

In each country, moderators trained in performing qualitative research conducted two FGDs in large cities in 2012. One of the authors (DD) debriefed the Polish and Swedish moderators before the discussion about background of the study and the question route. All participants gave written informed consent prior to the discussions. FGDs lasted for approximately 90

minutes and were conducted in the native language. All FGDs were audio taped and field notes were made during each discussion. At the end of the FGDs, all participants completed a short questionnaire on socio demographics and previous experiences with preventive measures.

The discussions were transcribed verbatim and identifiable data was removed. The entire Swedish and Polish transcripts, and the selected Dutch quotes were translated into English. A thematic analysis was performed [14, 16]. First, two authors (DD and IK) independently read all transcripts in-depth. Second, they created a provisional coding tree, based on the themes that emerged from the data. Third, they each identified and coded relevant text passages in one transcript per country and refined the coding tree. Perceived discrepancies between coders were discussed until consensus was reached and the coding tree was finalized. Fourth, one author (DD) coded the three remaining transcripts using the final coding tree (Additional file 3) and discussed her findings with IK. All transcripts and codes were imported into NVivo software (version 10, http://www.qsrinternational.com/) to enable systematic comparisons between different countries. We followed the COREQ-checklist when writing this paper [17].

RESULTS

In total, 41 people participated in six FGDs (Table 1). The median age ranged from 40 (Sweden) to 47 (the Netherlands). Approximately half of the participants were female. Lower and higher educated people participated in each FGD. The results are presented according to the themes that emerged from the data and were used in the final coding tree (Additional file 3). Representative quotations for each theme were selected to illustrate the results. If a quotation characterizes a minority opinion, it is indicated. The quotations are numbered; an additional label refers to the FGD ID.

Table 1 | Summary of participants' characteristics (n=41).

	Dutch participants[1] (n=17)	Polish participants[1] (n=12)	Swedish participants[1] (n=12)
Median age in years (range)	47 (22–77)	46 (19–61)	40 (21–80)
Female	8/17	6/12	6/12
Low educational level[2]	12/17	6/12	8/12
Having children	9/17	8/12	3/12
Belonging to risk group	8/17	2/12	2/12
If yes, seasonal influenza vaccine	7/8	0/2	1/2
Vaccinated against H1N1	9/17	0/12	10/12
Preventive measures against H1N1, other than vaccination	4/17	2/12	9/12
If yes, type of preventive measures[3]			
Hygiene[4]	4/4	2/2	6/9
Use nose-mouth mask	–	–	1/9
Avoid travelling abroad	–	–	1/9
Avoid crowded places	–	–	1/9
Not specified	–	–	3/9

Notes: [1] We conducted two focus group discussions in the Netherlands with nine and eight participants, respectively. In Poland and in Sweden six persons participated per discussion. [2]In all countries, high educational level was defined as tertiary education; all other educational levels were defined as 'low'. [3]Some participants stated that they applied multiple measures. [4] Hygiene includes washing hands more often, use hand sanitizer, cleaning desktop more often, etc.

1. Pandemic outbreak

Participants of all countries argued that their degree of concern for a new disease would depend on the mode and speed of transmission. They also would want to know the consequences of a disease, especially if potentially fatal, before deciding to take preventive measures or not. Often, comparisons with previous communicable disease outbreaks were made:

'I think it is all about this danger. If there is to be a new swine flu, maybe you will not actually get vaccinated, because you think it's not that dangerous. But if there is an Ebola epidemic. . .' (Q1SE1).

Dutch participants discussed that there would be no immunity for an outbreak with a new virus, thus resulting in uncertainty regarding the course of the disease. Swedish participants reasoned that they would experience the threat of a new disease as severe because they live in such a safe country:

'We do not have many other dangerous things to compare [the disease] with, so small things become dangerous to us' (Q2SE1).

Participants stated that they would weigh the threat of a new disease within the context of their own health status. Except for those who belong to a risk group (diabetes, asthma), most participants expected a low infection risk, e.g. thanks to healthy eating and living, and good personal hygiene. Participants considered information about the health status of infected people important, especially when also young and healthy people are infected, since this might change their perceived susceptibility. In addition, the proximity of cases was considered important; the closer the physical distance or the emotional relationship with a case, the higher the level of perceived susceptibility. However, it was remarked that proximity would be especially important in case of a severe disease. If relatives fell ill and the disease was not severe, participants did not intend to take safety measures. Most participants considered it wise to avoid visiting countries whit many people infected. Dutch and Swedish participants expressed their worry that communicable diseases might spread more rapidly nowadays:

'And now we move so incredibly easily: we fly and sail across the world. It can spread so easily' (Q3SE1).

In one of the Dutch groups it was put forward that lack of herd immunity due to large groups of unvaccinated people [e.g. in the so called 'Bible belt' in the Netherlands] might increase the risk of getting the disease. Polish participants did not discuss perceived seriousness of the disease and perceived susceptibility to the disease frequently.

2. Vaccination

Across all countries, preventing the disease or reducing the severity of its symptoms was considered the most important benefit of vaccines. Participants stated that the need for an effective vaccine would be higher when a disease was perceived as more serious. Anticipated regret made Dutch participants less hesitant to get vaccinated (Q4), even if the effectiveness of the vaccine was unclear.

'Doing something is better than doing nothing. . . if something is available you need to try it' (Q4NL1).

Several Swedes expressed that getting vaccinated would not only be beneficial for themselves but would also prevent them from infecting other people. Polish participants however were sceptical about the availability and quality of the vaccines in their country:

'Still you have to consider the fact that even if you got a loan just to get vaccinated, there probably wouldn't be any vaccines available in Poland, as usual' (Q5PL2).

[Participant 1] 'Don't you get this impression, which I have, that they ['like France, Denmark, the West'] get better vaccines while we get just the worst sort?'. [Participant 2] 'Yes'. [Participant 3] 'They get the first grade while we get the fourth grade. We import it, so we get the leftovers' (Q6PL2).

Several participants were opposed to vaccination in general as they believed it is healthiest if a body clears the virus without taking drugs:

'Why protect yourself against everything, while, in my opinion, it's more beneficial to have the disease and fight it yourself' (Q7NL2).

The most common view however was to weigh potential benefits and barriers of the vaccine against the threat of the disease:

'What's worse? Getting very sick and dying, or suffering from side effects? You do have to make a choice' (Q8NL1).

In general, the more severe the disease was seen, the less important the barriers to vaccination were considered:

'I got vaccinated against flu once and it's taken a great toll on me. I had high fever and headaches for three days. . . though if my life was in danger. . . I'd get vaccinated' (Q9PL1).

Participants displayed concerns about the safety of the vaccine due to the acute situation of a new pandemic, and limited time to produce and test vaccines and their safety:

'It will probably go damn fast, and they will not have time to test it. And therefore we will have no clue about the possible side effects' (Q10SE1).

Dutch participants expected their government to only introduce vaccines if they were considered safe:

'I do not expect the government to introduce a vaccine if they do not trust it themselves, or if they do not have insights into the side effects' (Q11NL1).

Across all countries, costs of the vaccine appeared to be a strong motivator in favour or against vaccination. Some participants stated that the price of a vaccine should not matter because life is precious, while others suggested that a vaccine should be available for the whole population and thus provided for free (e.g. by the government) (Q12, Q13). However, Polish participants did not believe that providing a vaccine for free would happen in their country (Q14).

[Participant 1] 'I think that if we have something so dignifiedly called public health, we should make it free'. [Participant 2] 'I think so too'. [Participant 3] 'I think so too, not everyone can spare a hundred and fifty Swedish Kronor' (Q12SE1).

'The costs are, in my opinion, the responsibility of the government. The government should simply protect its people, without putting a price tag on it' (Q13NL1).

'We would see the Prime Minster or the Minster of Health, who would tell us that the Polish government has decided to buy this vaccine and to provide it to us for free [laughs], which we wouldn't believe' (Q14PL1).

3. Social influences

Across all FGDs, it was expected that the new disease and vaccination would be discussed extensively by traditional mass media and on the Internet. Participants also frequently mentioned that this kind of information should be approached critically and that the source of information would really matter. One similar message disseminated across all media would be considered as more reliable. In general, Swedish participants were most trustful towards the national media (Q15), although some were critical (Q16).

'We have serious news reports, what is said in the news that is true' (Q15SE1).

'I was very sceptical of all the media pressure, and how they pointed out people, saying they were not showing solidarity because they did not get vaccinated, that everyone has to do it' (Q16SE1).

It was stated in the Dutch discussions that to prevent public panic, the government is expected to spread complete and trustworthy information as early in the pandemic as possible. Both Poles and Swedes agreed that in case of a pandemic a representative expert needs to step forward with the truth regarding the disease and vaccination (Q17), although Poles questioned the availability of such a person with that level of power and knowledge (Q18).

'You have to hope and believe that the medical community will step forward and honestly declare that it is safe, or that forty percent can experience side effects' (Q17SE2).

[Participant 1] 'So it's reliable knowledge provided by someone who's competent'. [Participant 2] 'We have no such authorities'. [Participant 3] 'All the good professors, specialists moved to the West' (Q18PL2).

Swedish participants considered the advice of relatives helpful in the decision about vaccination, while Polish participants did not:

'We [Polish people] might discuss it with someone, but everyone makes such decisions on their own' (Q19PL1).

The majority of Dutch participants expected to be personally invited by their general practitioner should they belong to a target group for vaccination:

'I think that if I belonged to the target group, I would be invited automatically by my general practitioner' (Q20NL1).

In the Dutch and Swedish discussions participants suggested that they would contact people who already have been vaccinated, to learn from their experience. However, participants were unsure if those opinions would alter their decision. Participants mentioned that seeing friends or family suffer from the disease, would make them feel not only more susceptible to the disease, but also more willing to get vaccinated:

'I think that if someone close to me or an acquaintance of mine died of this disease, then it would decidedly make me get vaccinated faster' (Q21PL2).

The vaccination decision of quite some Dutch and Swedish participants would be influenced by the vaccination behaviour of the majority of their peers. Although some stated that revising their opinion would depend on the number of and their relation with vaccinated peers:

'If everyone in your vicinity gets vaccinated, it is clear that it will affect my decision. Then I will begin to wonder: should I really ignore this?' (Q22SE2).

4. Population characteristics

It was mentioned by Swedish participants that during the H1N1 pandemic applying preventive measures was an automatic response to the government's call to get vaccinated, and that it was an exception if one did not get vaccinated. They concluded that they were a generally risk aware, obedient, and very serious and equality focused population:

'It was true that the authorities stepped forward and told everyone to get vaccinated. It was almost a command. You felt a bit guilty if you did not do it, I think' (Q23SE2).

'We [the Swedes] are quick to agree with each other, and then we go home and grumble a little on our own. It's a mentality. We are such herd animals; we do what everyone else does' (Q24SE2).

'We want to do the right thing. When you are sitting in your car, you should wear your seat belt. And if someone says that we will all get sick, so now you should take a vaccine, then I take that vaccine' (Q25SE2).

Polish participants mentioned being sceptical and reluctant regarding vaccines and to be somewhat lacking in trust in doctors and the production process of vaccines:

'I think they ['conscious societies', Norway is given as an example] would obediently arrive for the vaccination, and they wouldn't hesitate. Whereas here [in Poland], people would start to speculate just like we're speculating now. Should we do it, or maybe it's not worth it, or maybe the devil's in the detail' (Q26PL1).

'Abroad everyone trusts doctors. It's scary in a way. They have a completely different attitude to doctors' (Q27PL1).

5. Prior contact with similar diseases or vaccinations

Participants frequently referred to their experience with the Influenza A/H1N1 2009 pandemic throughout the discussions, also without the moderators introducing this topic. Dutch and Swedish participants stated that due to their experience with the H1N1 pandemic they would perceive any new disease as less serious (Q28). Additionally, Swedish participants were sceptical regarding the safety of vaccines because of the debate concerning narcolepsy as a side effect of the 2009 pandemic vaccine (Q29).

'The risk when a new one [a new outbreak] comes. . . Many may think that it is exaggerated, like the swine flu was' (Q28SE2).

'But this [debating about the safety of vaccines] is a new phenomenon. Before the swine flu came, we had never had this debate. People have been vaccinated for who knows how many years' (Q29SE2).

Participating Poles reflected positively on their governments' decision not to buy the pandemic vaccine:

'Well, there was this propaganda to get vaccinated. Of course there was! There was propaganda all around the world. But it was limited in Poland and that's good, because it turned out we were the only country in Europe that didn't lose face then' (Q30PL1).

In addition, experiences with a previous pandemic or seasonal flu may affect choices to get vaccinated for a new disease, in these cases positively:

'My neighbour, a healthy boy of 13 years old, died of it [the Hong Kong flu]. In my opinion, it is not relevant that there is a chance that the shot doesn't work or that the outbreak will not end in an epidemic…if there are no horrible stories about it [the vaccine], I'll take the shot' (Q31NL1).

'I had severe flu complications several years ago. I ended up in the hospital, it was horrible. A disease like that makes you change the way you think [regarding vaccination]' (Q32PL2).

6. Health authorities

Many participants put forward that doctors do not always agree on the use of preventive measures during pandemics. National Public Health Institutes were frequently mentioned as being trust- worthy and reliable sources of information during Dutch (Q33) and Swedish discussions, but not mentioned in the Polish discussions. Instead, participants complained about the status of the public health system in Poland (Q34).

'If the outbreak is as severe as you describe just now, the RIVM [National Institute for Public Health and the Environment in the Netherlands] needs to play an active role, and inform us, instead of us being dependent on subjective information' (Q33NL2).

'Prevention is more common there [in the West of Europe]. Maybe they feel protected by the state more. We don't have that comfort' (Q34PL1).

Polish participants were sceptical and distrustful when discussing their government, while the Swedish groups frequently discussed their trust in government and the tendency to obey the government, in spite of the decrease in trust since the H1N1 pandemic (Q35, Q36). They also mentioned the lack of trust in the government elsewhere.

'You were really taken by surprise: My God, the state has given us something that was not good. You're not used to it, after all' (Q35SE2).

'During the swine flu days, the initial stand on the vaccine was: Everyone should take it, and it's safe, we're all going to die, so you have to get vaccinated. And then suddenly it changed: No, no, Sweden has signed an agreement about this vaccine. We [the Swedish nation] had to buy it, which meant that they [the pharmaceutical companies] wanted to sell it, and then it turned out that it had not actually been tested. I think that is crazy' (Q36SE2).

All groups discussed that people would want to make money on new vaccines. These expectations influenced participants' opinion on getting vaccinated negatively, although several Dutch and Swedish participants tended to trust and defend their governments:

'I cannot keep on being so terribly sceptical. . . I have decided that there are some government bodies that you trust. Otherwise I would probably feel that there is no point that they exist' (Q37SE1).

In both Dutch discussions, the advantages of international cooperation regarding the outbreak and vaccination were put forward.

DISCUSSION

We explored public opinion and attitudes regarding vaccination during future pandemics and possible national differences by conducting FGDs in three European countries: the Netherlands, Poland and Sweden. Participants stated that they would base their vaccination decision on trade-offs between perceived benefits and barriers of the vaccine, also taking into account the seriousness of a new pandemic outbreak. Except for those who belong to a risk group, most participants in the present study expected a low infection risk, resulting in a lower willingness to get vaccinated. A questionnaire study on seasonal Influenza vaccination coverage and reasons to refrain among high-risk persons in four European countries, including Poland and Sweden [18] showed that individuals did not perceive themselves as susceptible to seasonal Influenza either. During future outbreaks, it is therefore necessary to provide the public with information regarding the health status of first cases, especially when also young and healthy people are infected, with information about the general level of susceptibility and a specification of which groups are considered vulnerable and are thus being targeted for vaccination. The displayed concerns regarding the safety of newly developed vaccines were also observed in a Canadian focus group study [4]; people were hesitant to accept vaccines during future pandemics due to the perceived uncertainties considering novel vaccines and seriousness of disease.

Importantly, some differences between European countries were observed that have implications for outbreak preparedness. We did observe differences in adherence to social norms and rules. Whereas Swedish participants displayed a tendency to do the right thing and to get vaccinated to protect others, this appeared much less prominent in the Dutch and Polish participants. In countries where there is a culture to follow social norms, such as in Sweden, communication might focus more on the social norm, e.g. by providing normative information, both descriptive (perception of the proportion of people opting for vaccination) and injunctive (perception of what is approved or disapproved by others) [19–

21]. Trust in health authorities (or lack thereof) has implications for outbreak planning too. Dutch and Swedish participants displayed more trust in both health professionals as well as in national governments than Polish participants. This was also observed in a survey during the Influenza A/H1N1 2009 pandemic [22]. These different levels of trust have implications for the promotion of and response to public health messages from national governments and their public health agencies [23–26]. The lack of trust in e.g. statements issued by the national government can be problematic, since this has been linked to a reduction in vaccination behaviour [23, 25, 27, 28].

It is important to build trust in the pre-outbreak phase, maintain trust during outbreaks and, if necessary, restore or further develop trust after the pandemic ends [29,30]. To do so, reliable and trusted local representatives of the medical community need to communicate clear public health messages regarding the new outbreak and preventive measures. Swedish participants indicated that their experiences during the Influenza A/H1N1 2009 pandemic would reduce their tendency to accept vaccination advice. These discussions may be rooted in the Swedish government having signed a contract with a pharmaceutical company to buy pandemic Influenza A/H1N1 vaccines years before the outbreak [31] and the high incidence rates of narcolepsy following the Influenza A/H1N1 2009 pandemic, suggesting an association with vaccination [32, 33]. The seasonal Influenza vaccination coverage in Sweden decreased since the Influenza A/H1N1 2009 pandemic; it was 65.8% in 2008–2009 but decreased to 44.3% in 2012–2013 [34]. Combined with the Polish participants being proud that their Minister of Health had not bought vaccines during the Influenza A/H1N1 2009 pandemic, these findings confirm what Börjesson et al. concluded in 2013; previous experiences with outbreak situations play a crucial role in public opinions and future behaviour. Our study highlights that outbreak experiences differ between countries in many dimensions: with regard to cultural differences, with regard to government policies, and with regard to vaccination side effects (narcolepsy in Sweden). These differences stress the need to adapt communication strategies to local circumstances.

Although efforts to include individuals of different gender, age and educational level were successful, there might still be responder bias; individuals who participated might be particularly interested in the topic. This paper provides an illustration of opinions and attitudes regarding future pandemics and vaccination among members of the general public in three different European countries. Future research could also focus on opinions and attitudes of health care workers across European countries because of the example they represent for public opinion. Vaccination history as well as intentions of the general public to be vaccinated are positively associated with recommendations by health care workers to do so [35]. Conclusions drawn from this study should be considered with some caution as the findings are based on a small number of individuals, and may therefore not be generalizable to populations at large. In addition, we cannot exclude the possibility that some of the

observed differences relate to individual differences rather than to differences between countries. However, as there is hardly any research examining differences in opinions and attitudes regarding pandemics and vaccination across Europe, our results can be seen as a first step in this process.

REFERENCES

1. Evolution of a pandemic: A(H1N1) 2009, April 2009–August 2010. World Health Organization (WHO); 2013. Available from: http://apps.who.int/iris/bitstream/10665/78414/1/9789241503051eng.pdf?ua=1
2. Briand S, Bertherat E, Cox P, et al. The international Ebola emergency. N Engl J Med. 2014;371:1180–3.
3. Nguyen-Van-Tam JS, Hampson AW. The epidemiology and clinical impact of pandemic influenza. Vaccine. 2003;21:1762–8.
4. Henrich N, Holmes B. The public's acceptance of novel vaccines during a pandemic: a focus group study and its application to influenza H1N1. Emerg Health Threats. J 2009;2:e8.
5. Lipsitch M, Finelli L, Heffernan RT, et al. Improving the evidence base for decision making during a pandemic: the example of 2009 influenza A/H1N1. Biosecur Bioterror. 2011;9:89–115.
6. Morrison LG, Yardley L. What infection control measures will people carry out to reduce transmission of pandemic influenza? A focus group study. BMC Public Health. 2009;9:258.
7. Bish A, Yardley L, Nicoll A, et al. Factors associated with uptake of vaccination against pandemic influenza: a systematic review. Vaccine. 2011;29:6472–84.
8. Nguyen T, Henningsen KH, Brehaut JC, et al. Acceptance of a pandemic influenza vaccine: a systematic review of surveys of the general public. Infect Drug Resist. 2011;4:197–207.
9. Brien S, Kwong JC, Buckeridge DL. The determinants of 2009 pandemic A/H1N1 influenza vaccination: a systematic review. Vaccine. 2012;30:1255–64.
10. Determann D, Korfage IJ, Lambooij MS, et al. Acceptance of vaccinations in pandemic outbreaks: a discrete choice experiment. PLoS ONE. 2014;9:e102505.
11. Davis MD, Stephenson N, Lohm D, et al. Beyond resistance: social factors in the general public response to pandemic influenza. BMC Public Health. 2015;15:436.
12. Teasdale E, Yardley L. Understanding responses to government health recommendations: public perceptions of government advice for managing the H1N1 (swine flu) influenza pandemic. Patient Educ Couns. 2011;85:413–8.
13. Powell RA, Single HM. Focus groups. Int J Qual Health Care. 1996;8:499–504.
14. Morgan DL, Krueger RA. The focus group kit. London: SAGE Publications, Inc.;1997.
15. Kitzinger J. Qualitative research. Introducing focus groups. BMJ. 1995;311:299–302.
16. Pope C, Ziebland S, Mays N. Qualitative research in health care. Analysing qualitative data. BMJ. 2000;320:114–6.
17. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care. 2007;19:349–57.
18. Kroneman M, van Essen GA, Paget WJ. Influenza vaccination coverage and reasons to refrain among high-risk persons in four European countries. Vaccine. 2006;24:622–8.
19. Schultz PW, Nolan JM, Cialdini RB, et al. The constructive, destructive, and reconstructive power of social norms. Psychol Sci. 2007;18:429–34.
20. Nolan JM, Schultz PW, Cialdini RB, et al. Normative social influence is underdetected. Pers Soc Psychol Bull. 2008;34:913–23.
21. Nolan JM, Schultz PW, Knowles ES. Using public service announcements to change behavior: no more money and oil down the drain. J Appl Soc Psychol. 2009;39:1035-56.
22. Influenza H1N1 Analytical Report – Eurobarometer on Influenza H1N1. The Gallup Organization; 2010. Available from: http://ec.europa.eu/public opinion/flash/fl 287 en.pdf.
23. Van der Weerd W, Timmermans DR, Beaujean DJ, et al. Monitoring the level of government trust, risk perception and intention of the general public to adopt protective measures during the influenza A (H1N1) pandemic in the Netherlands. BMC Public Health. 2011;11:575.
24. Siegrist M, Earle TC, Gutscher H. Test of a trust and confidence model in the applied context of electromagnetic field (EMF) risks. Risk Anal. 2003;23:705–16.

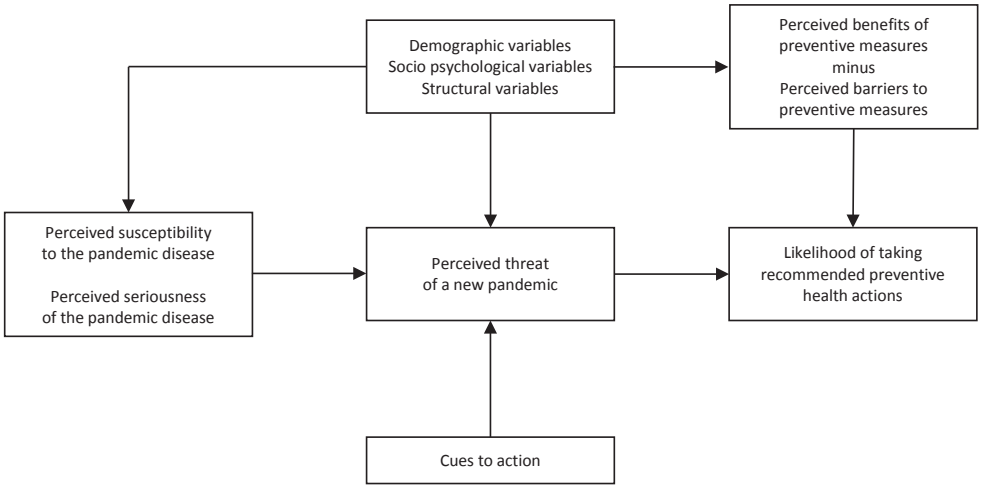
25. Gilles I, Bangerter A, Clemence A, et al. Trust in medical organizations predicts pandemic (H1N1) 2009 vaccination behavior and perceived efficacy of protection measures in the Swiss public. Eur J Epidemiol. 2011;26:203–10.
26. Larson HJ, Heymann DL. Public health response to influenza A(H1N1) as an opportunity to build public trust. JAMA. 2010;303:271–2.
27. Allen JD, Othus MK, Shelton RC, et al. Parental decision making about the HPV vaccine. Cancer Epidemiol Biomarkers Prev. 2010;19:2187–98.
28. Ronnerstrand B. Social capital and immunisation against the 2009 A(H1N1) pandemic in Sweden. Scand J Public Health. 2013;41:853–9.
29. Paek HJ, Hilyard K, Freimuth VS, et al. Public support for government actions during a flu pandemic: lessons learned from a statewide survey. Health Promot Pract. 2008;9:60S–72S.
30. Gesser-Edelsburg A, Mordini E, James JJ, et al. Risk communication recommendations and implementation during emerging infectious diseases: a case study of the 2009 H1N1 influenza pandemic. Disaster Med Public Health Prep. 2014:1–12.
31. Widgren K, Magnusson M, Hagstam P, et al. Prevailing effectiveness of the 2009 influenza A(H1N1) pdm09 vaccine during the 2010/11 season in Sweden. Euro Surveill. 2013;18:20447.
32. Barker CI, Snape MD. Pandemic influenza A H1N1 vaccines and narcolepsy: vaccine safety surveillance in action. Lancet Infect Dis. 2014;14:227–38.
33. Wijnans L, Lecomte C, de Vries C, et al. The incidence of narcolepsy in Europe: before, during, and after the influenza A(H1N1)pdm09 pandemic and vaccination campaigns. Vaccine. 2013;31:1246–54.
34. Seasonal influenza vaccination programme country profile: Sweden (2012–13 Season). European Centre for Disease Prevention and Control (ECDC); 2014. Available from: http://ecdc.europa.eu/en/publications/Report%20Assets/seasonal- vaccination-coverage-in-europe-2012-13/Seasonal-Influenza-Vaccination- Programme-Country-Profile-Sweden.pdf.
35. Schwarzinger M, Flicoteaux R, Cortarenoda S, et al. Low acceptability of A/H1N1 pandemic vaccination in French adult population: did public health policy fuel public dissonance? PLoS ONE. 2010;5:e10199.

ADDITIONAL FILE 1: BEHAVIOURAL MODELS

We selected the Health Belief Model (HBM) [1] to create the focus group discussion guide because of its large empirically tested ability to explain and predict intention of complying with preventive medical care recommendations [2, 3].

The basic idea underlying the HBM (Figure 1) is that people react to a perceived threat, in this case a new pandemic outbreak, by performing some action. The perceived threat depends on both the perceived susceptibility to the disease (a person's perception of the chance that he/she will contract the disease) and the perceived severity of the disease (a person's belief on how serious contracting the disease would be for him/her, both medical and social consequences). According to the HBM model, people weigh this perceived threat to the perceived benefits of actions (beliefs regarding the effectiveness of preventive measures in reducing the disease threat) and to the perceived barriers of actions (potential negative aspects of the preventive measures). The HBM posits that prevention is more likely if there is a high threat, if people believe an available action will reduce their susceptibility or severity of the condition and if the barriers to actions are outweighed by the benefits. If perceived severity and susceptibility are low, people will not perceive the disease as threatening and will consequently not be inclined to act. Besides weighing the perceived threat to the perceived efficacy of the vaccination program, the HBM states that 'cues to action' (strategies to activate readiness) are affecting someone's intention as well. These cues can either be internal (e.g. symptoms) or external (e.g. mass media campaigns, advice from others, illness of family member of friend, newspaper or magazine article, reminder)). Furthermore, according to the HBM the decisions people make are also dependent on certain variables that can be classified as: demographic (age, sex, race, ethnicity, etc.) and social psychological (personality, social class, peer and reference group behaviour, etc.) and structural (knowledge about the disease, prior contact with the disease).

Figure 1 | Overview of Health Belief Model [1], adapted for the current study to new pandemic outbreaks.



We also incorporated two concepts from other theories: 1) 'self-efficacy' (a person's level of confidence in his/her ability to perform the preventive measure) from Protection Motivation Theory [4] and Bandura's Social Cognitive Theory [5], and 2) emotions and affect which play a role in decision making [6, 7], also when deciding on prevention for infectious disease outbreaks [8].

REFERENCES

1. Janz NK, Becker MH. The Health Belief Model: a decade later. Health Educ Q. 1984;11:1-47.
2. Coe AB, Gatewood SB, Moczygemba LR, et al. The use of the health belief model to assess predictors of intent to receive the novel (2009) H1N1 influenza vaccine. Inov Pharm. 2012;3:1-11.
3. Teitler-Regev S, Shahrabani S, Benzion U. Factors Affecting Intention among Students to Be Vaccinated against A/H1N1 Influenza: A Health Belief Model Approach. Adv Prev Med. 2011;2011:353207.
4. Rogers RW. A protection motivation theory of fear appeals and attitude change. JRL. 1975;91:93-114.
5. Bandura A. Social foundations of thought and action: a social cognitive theory: Englewood Cilffs: Prentice-Hall; 1986.
6. Loewenstein GF, Weber EU, Hsee CK, et al. Risk as feelings. Psychol Bull. 2001;127:267-86.
7. Slovic P, Finucane ML, Peters E, et al. Risk as analysis and risk as feelings: some thoughts about affect, reason, risk, and rationality. Risk Anal. 2004;24:311-22.
8. Chapman GB, Coups EJ. Emotions and preventive health behavior: Worry, regret, and influenza vaccination. Health Psychol. 2006;25:82-90.

ADDITIONAL FILE 2: FOCUS GROUP QUESTION ROUTE

Request for the discussion leader:

Ask the participants to sign and hand in the informed consent form.
Ask the participants to write their first names on the nameplates.

Topic list, to be discussed during the focus group discussion:

Welcome

- Welcome to this focus group discussion about vaccinations. Thanks for coming!
- My name is … and I will lead the discussion today. Next to me is …, a researcher at … and she will assist me today.

Introduction by researcher

- Why do we do this study? It could be that there will be an outbreak of a new contagious disease in the future. Then, it's important to know whether people are willing to get vaccinated against this disease. It's also important to know what reasons people may have to get vaccinated or not.
- Goal of today: discuss people's opinions about vaccination and whether they are willing to get vaccinated against a new contagious disease. We are interested in your opinions as … citizens. So for tonight, imagine that you are living in … again.

Practical

- Rules: everybody is free to give their opinion. There is no need to reach consensus. My role is limited; I will only lead the discussion, you will discuss as a group.
- I would like to ask you to keep everything that will be said today confidential.
- The discussion will be recorded by using a voice recorder, so that we know what has been said literally later. When analysing the discussion, this will be done anonymously, no names will be mentioned.
- What does this meeting look like? The discussion will last for approximately 2 hours. Halfway, there will be a short break from 5 to 10 minutes. You'll first discuss the new contagious disease. Thereafter you will discuss about a vaccination against this new disease.
- Are there any questions before we'll start?

[Turn on the voice recorder]

We'll start with a short introduction:

- What is your name and the reason that you are participating today?

1) Is there anyone who can remember it: what was the last vaccination you got?
    a. Which vaccination was it? When was it?
    b. What were reasons for you to get vaccinated then?

From now on, we'll focus on vaccinations against new contagious diseases.

Imagine a new contagious disease emerged abroad. People already died of the disease there. People got sick in … as well, but no one died yet.

2) What would you like to know about this disease? And why?
3) Are you actively going to search for information? Where are you looking for this information?
4) Which media do you use to search for information about this new disease?
5) How would you like to get informed about the new disease?
6) Who needs to give this information?

7) Are you afraid to get sick? Why do you? Why don't you?
8) Will you look for preventive measures you can take against the disease? Which measures?

Imagine a vaccination has to be produced against this new disease.

9) Some people only get vaccinated if they think the vaccination is 'good'. What do you think is a good vaccine/ when do you think a vaccine is 'good'? What attributes does the vaccine need to have?

10) How much would you like to pay for the vaccination?
11) What role does the price play in your decision to get vaccinated or not?

We'll have a short break of 5 – 10 minutes. We'll start again at…. pm.

After the break:

Now imagine that the vaccine has been produced and you are eligible to get a vaccine as well.

12) Will you be informed of this? If yes, by whom will you be informed?
13) Will you be going to ask for advice before deciding about uptake of the vaccine? If yes, whom would you ask? And what do you do with this advice?
14) Do you take the decision to get vaccinated only based on facts? If no, what is also important?
15) Do your previous experiences with vaccinations influence the decision you have to make now?

Imagine a person that is really important for you has already been vaccinated.

16) Does this affect your own decision, and if yes, how?

Imagine, you've searched for information about the disease and the vaccine and you discussed about the vaccine with your friends and family. If you have decided to get vaccinated, you have an opportunity to do this tomorrow.

17) Is there anything that can stop you from getting vaccinated? If yes, what?
18) What practical matters are important now?

We will now discuss a couple of situations.

Scenario 1: All over the world more people get sick and more people are dying from the new disease. Also in …, more people are getting sick. The first … person died.

19) The situation has changed. What does this change mean for you?

Scenario 2: The whole world is under the spell of the new disease. More people died in … as well, because of the disease. Also a healthy … girl died.

20) The situation has changed. What does this change mean for you?

Scenario 3: Now, also someone in your close circle got infected with the new disease. It is unclear if he or she will survive. This person wasn't vaccinated against the new disease.

21) The situation has changed. What does this change for you?
22) What if this person has been vaccinated, but still contracted the disease?

All the previous questions were about a vaccination for you. You now have to decide if you want your child to get vaccinated against the new disease. If you do not have children, please imagine for now that you have.

23) What (other) reasons are important now, when you need to decide if your child gets the vaccination or not?

Until now, we discussed about what you would do if a new contagious disease will break out. Now, think back at the Mexican flu (or swine flu, or H1N1) outbreak from a couple of years ago. Remember, we are talking about the … situation.

24) Were you worried to get sick then? Why? Why not? Did anyone in your close circle got sick of H1N1?
25) What did you think of the information available about the H1N1 and the precautions that you could take against the H1N1?
26) What did you do to avoid getting the H1N1?
27) The last question: Do you think there are differences between … and other European countries in how people react to a new disease and in reasons to get vaccinated or not? If yes, what are those differences? Also think about differences in which things are organized.

Requests for discussion leader:

Ask the participants if they have anything to add to the things that are said today?
Ask the participants if they have other questions or remarks about the discussion?

[Turn off the voice recorder]

End the conversation. Thank the participants for their participation.

Ask the participants to fill in the questionnaire.

Distribute the gift vouchers. Ask the participants to sign after receiving the voucher.

ADDITIONAL FILE 3: FINAL CODING TREE

1. Pandemic outbreak
    a. Perceived seriousness of contracting the disease
    b. Perceived susceptibility of contracting the disease

2. Vaccination
    a. Perceived benefits of vaccinations
    b. Perceived barriers to vaccinations

3. Social influences
    a. Mass media
    b. Advice from others
    c. Illness of friends and family
    d. Peer behaviour

4. Population characteristics

5. Prior contact with similar diseases and vaccinations

6. Health authorities (health care practitioners, researchers, pharmaceutical companies, Institutes of Public Health, Ministry of Health, International Organizations, such as WHO).

4

Public preferences for vaccination programmes during pandemics caused by pathogens transmitted through respiratory droplets - a discrete choice experiment in four European countries, 2013

Domino Determann
Ida J. Korfage
Angela Fagerlin
Ewout W. Steyerberg
Michiel C. Bliemer
Helene A. Voeten
Jan Hendrik Richardus
Mattijs S. Lambooij
Esther W. de Bekker-Grob

Eurosurveillance. 2016;21(22):pii=30247
doi:10.2807/1560-7917.ES.2016.21.22.30247

ABSTRACT

This study aims to quantify and compare preferences of citizens from different European countries for vaccination programme characteristics during pandemics, caused by pathogens which are transmitted through respiratory droplets. Internet panel members, nationally representative based on age, sex, educational level and region, of four European Union Member States (Netherlands, Poland, Spain, and Sweden, n = 2,068) completed an online discrete choice experiment. These countries, from different geographical areas of Europe, were chosen because of the availability of high-quality Internet panels and because of the cooperation between members of the project entitled Effective Communication in Outbreak Management: development of an evidence-based tool for Europe (ECOM). Data were analysed using panel latent class regression models. In the case of a severe pandemic scenario, vaccine effectiveness was the most important characteristic determining vaccination preference in all countries, followed by the body that advises on vaccination. In Sweden, the advice of family and/or friends and the advice of physicians strongly affected vaccine preferences, in contrast to Poland and Spain, where the advice of (international) health authorities was more decisive. Irrespective of pandemic scenario or vaccination programme characteristics, the predicted vaccination uptakes were lowest in Sweden, and highest in Poland. To increase vaccination uptake during future pandemics, the responsible authorities should align with other important stakeholders in the country and communicate in a coordinated manner.

INTRODUCTION

In the past 100 years, there have been several large-scale influenza outbreaks with worldwide impact. These include the 1918 influenza A(H1N1) pandemic that caused between 50 and 100 million deaths particularly in many healthy young adults [1], and more recently the 2009 influenza A(H1N1)pdm09 pandemic [2]. Though characteristics (such as clinical attack rates and pathogenicity) and occurrence of a next influenza pandemic are unpredictable, experts agree there will be future influenza pandemics [2-5].

The World Health Organisation (WHO) urged countries to develop or update national influenza preparedness plans in response to the avian influenza A(H5N1) pandemic threat in 2005 [6]. Such plans subsequently needed to be improved taking into account the lessons learnt from the response to the influenza A(H1N1)pdm09 pandemic [4,7,8]. In addition, countries could learn from each other by sharing information and best practices [9].

Preventive measures are very important in limiting the spread of an influenza pandemic [10-12] and if available, vaccination constitutes the control cornerstone [13,14]. The success of mitigating influenza pandemics depends on many factors, including national public health policies and the availability of vaccines, vaccine effectiveness, and the public's willingness to get vaccinated. Unfortunately, vaccination coverage has proven to be (too) low across Europe during the influenza A(H1N1)pdm09 pandemic. Vaccination coverage among the general public of the European Union, Norway and Iceland, varied between countries from 0.4% to 59% [15].

Countries within Europe differ from each other with regard to languages, cultures, public trust in health authorities, health system infrastructures, and public health capabilities and capacities. Research has shown that implementing international guidelines at the local level can be a complex process [16]. Having insights into country-specific reasons to accept or decline pandemic influenza vaccination can facilitate the adaptation of preparedness plans, including vaccination strategies, to the local situation [17].

Thus far, only a limited number of reports have focused on the comparison of pandemic influenza vaccination preferences between people of different European countries [18,19], and formal quantitative techniques such as discrete choice experiments (DCEs) [20,21] have not yet been used. The primary aim of this study was to quantify and compare the preferences of European citizens for vaccination programmes for future pandemics. Although we focus on influenza pandemics, we quantified vaccination programme preferences for any emerging or re-emerging large-scale infectious disease outbreak that spreads through respiratory droplets. Our findings might therefore also be applicable to other respiratory infections than influenza, such as, for example, severe acute respiratory syndrome (SARS)-

coronavirus (CoV) or Middle East Respiratory Syndrome (MERS)-CoV, should vaccines be available for these viruses in the future. A secondary aim was to calculate the expected uptake of vaccination under different pandemic scenarios. The approach and results might help health policymakers to improve pandemic preparedness plans and communication strategies, in order to make future vaccination programmes more successful.

METHODS

Study population

We surveyed a representative sample of the general public (age 18 years and over) of countries from different parts of Europe: eastern Europe (Poland), northern Europe (Sweden), southern Europe (Spain) and western Europe (Netherlands). These countries were chosen because of the availability of high-quality Internet panels (i.e. panels that are ISO certified and/or follow international quality standards for market research) and also because of the cooperation between project members of different work packages within the Effective Communication in Outbreak Management: development of an evidence-based tool for Europe (ECOM) project (www.ecomeu.info). The public health policies of the four included countries with respect to seasonal influenza and influenza A(H1N1)pdm09 are described in Table 1.

Discrete choice experiments

A DCE is a survey-based stated-preference methodology that originates in mathematical psychology [22]. The method has been increasingly used in healthcare, whereby the number of published DCEs has increased from a mean of three per year in the period from 1990 to 2000 to 45 per year between 2009 and 2012 [23]. In a DCE, the relative importance of characteristics (i.e. attributes) of a certain product or intervention is assessed by presenting a series of choice sets to respondents [20,21]. In each choice set, respondents are asked to choose a preferred alternative from a set of two or more hypothetical product or intervention alternatives with systematically varying attribute levels [20,21].

Survey

The survey started with an explanation of the DCE exercise. Next, respondents were asked to imagine that a large-scale emerging infectious disease, that started abroad, had spread to the country they lived in. It was stated that the disease spreads through respiratory droplets, that it was vaccine-preventable, and that vaccines were available in their country. Respondents then completed a series of choice sets, followed by questions about socio-demographic characteristics (including previous vaccination experiences), and questions that assessed the perceived difficulty of the survey. The survey ended with an open question in which respondents were given the opportunity to comment on the survey.

In each choice set, a hypothetical pandemic scenario based on two disease variables (susceptibility to the disease (i.e. a number of 1,000 people will get sick) and severity of the disease (i.e. a number of the sick people will develop severe symptoms) was presented. Respondents were then asked to choose between three alternatives: no vaccination, vaccination A, and vaccination B. The vaccination was described by several attributes, and the presented levels differed systematically between vaccination A and vaccination B. In the following choice sets, both the pandemic scenario and the presented attribute levels for vaccination A and B differed. In order to select realistic, relevant and understandable attributes and attributes levels, we conducted a literature study, expert interviews, and focus group discussions. In addition, we closely cooperated with project members when selecting the attributes and levels. PubMed, Embase and Psychinfo were strategically searched for relevant research articles on vaccination preferences. Expert interviews (n = 9) were conducted with both national and international experts (physicians, researchers, policymakers) in the field of infectious diseases, vaccinations, preventive behaviour, and implementation of prevention. We conducted eight focus group discussions with representatives of the general population, of which four in the Netherlands, two in Poland, two with Spanish citizens during their temporary stay in the Netherlands, and two in Sweden. Eligible participants were recruited by research companies and via our network, using purposive sampling to ensure a diverse sample. The focus groups revealed that similar vaccination programme attributes and attribute levels could be included in the DCE for all countries (Table 2). It is not feasible to present a single respondent with all the possible combinations of the included attribute levels. We therefore generated a subset of 48 choice sets by minimizing the D-efficiency criterion using the software programme Ngene (ChoiceMetrics, version 1.1.1). The 48 choice sets were grouped in three different survey versions such that each block has (near) attribute level balance. Each respondent thus needed to answer 16 choice sets. For more information on this part of a discrete choice experiment, see e.g. Reed Johnson et al. [24].

Table 1 | Overview of seasonal influenza and influenza A(H1N1)pdm09 policies per country, Netherlands, Poland, Spain, and Sweden, 2009 and 2013.

Influenza type and respective policies	Netherlands	Poland	Spain	Sweden
Seasonal influenza [58]				
Groups recommended for vaccination during the 2012/13 influenza season	NA	Children and adolescents, aged ≥ 6 months – < 18 years	NA	NA
	Adults aged ≥ 60 years	Adults aged ≥ 55 years	Adults aged ≥ 65 years[1]	Adults aged ≥ 65 years
	Medical risk groups[2]	Medical risk groups[2]	Medical risk groups[2]	Medical risk groups[2]
	Pregnant women with medical conditions	All pregnant women	All pregnant women	Pregnant women in 2nd or 3rd trimester
	All HCWs	All HCWs	All HCWs	HCWs caring for persons who are severely immunocompromised
Payment scheme vaccine and administration during the 2012/13 influenza season	National health service Employer pays for HCWs	Payment scheme vaccine itself: out-of-pocket; some employers pay for HCWs; local government[3] Payment scheme administration: out-of-pocket; some employers pay for HCWs; local government[3]	Regional health service	Regional health service; out-of-pocket varies with regions[4] Employer pays for HCWs
Vaccination coverage during the 2012/13 influenza season	Overall adults aged ≥ 60 years: 67.8%	Overall adults aged ≥ 65 years: 7.4% HCWs: 9.5%	Overall adults aged ≥ 65 years: 57% HCWs: 22.9%	Overall adults aged ≥ 65 years: 44%
2009 influenza A(H1N1)pdm09 pandemic [19]				
Groups recommended for vaccination during the pandemic period	Children aged ≥ 6 months – 4 years, and household members of babies up to the age of 6 months	Poland did not implement a vaccination programme during the influenza A(H1N1)pdm09	NA	Recommended for all children aged ≥ 6 months – < 18 years
	Adults aged ≥ 60 years	NA	NA	Adults aged ≥ 18 years
	Medical risk groups[2]	NA	Medical risk groups[2]	Medical risk groups[2]
	Pregnant women in 2nd and 3rd trimester	NA	All pregnant women	All pregnant women
	HCWs with close contact with patients	NA	All HCWs	All HCWs

Vaccine brand	Pandemrix, Focetria	NA	Pandemrix, Focetria, and Panenza	Pandemrix
Vaccination sites	GPs, mass vaccination sites in community settings, Municipal Health Services (children and household contacts), and work environment	NA	GPs, hospital settings, and occupational health services	GPs, hospital settings, outpatient care clinics, occupational health services, mass vaccination sites
Payment scheme	Free of charge for all individuals recommended the vaccine	NA	Free of charge for all individuals recommended the vaccine	Free of charge for all individuals recommended the vaccine
Vaccination coverage during the pandemic period	Entire population: 30% Those at risk aged >6 months: 72% Pregnant women: 58% HCWs: 50%	NA	Entire population: 27.1% Those at risk aged >6 months: 23.7% Pregnant women: 9% HCWs: 11.6%	Entire population: 59%

Abbreviations: *GP* general practitioner; *HCW* healthcare worker; *NA* not applicable.
Notes: [1]Recommendation at the national level. However, 10 of 19 regions recommend vaccine for those ≥ 60 years. [2]Medical risk groups include e.g. patients with chronic pulmonary, cardiovascular and renal diseases, metabolic disorders, and immunosuppression due to disease or treatment (we refer to [1] for more details). [3]Local government reimbursement of cost of vaccine and administration for those ≥ 65 years of age. [4]In some regions, the vaccine is charged a symbolic amount (ca 10 euros) for vaccine and vaccination.

Table 2 | Attributes and attribute levels included in the survey investigating public preferences for vaccination programmes during pandemics caused by pathogens transmitted through respiratory droplets, Netherlands, Poland, Spain, and Sweden, 2013 (n = 7 attributes).

Pandemic scenario variables[1]	**Levels**
Susceptibility to the disease[2]	5%, 10%, 20%
Severity of the disease[3]	5%, 25%, 50%, 75%
Vaccination programme attributes[4]	**Levels**
Effectiveness of the vaccine	30%, 50%, 70%, 90%
Safety of the vaccine[5]	Unknown, expected to be safe (reference level)
	Unknown, no experience with similar vaccines yet
Advice regarding the vaccine	Family and/or friends recommend vaccination (reference level)
	Family and/or friends discourage vaccination
	Your doctor recommends vaccination
	Your doctor discourages vaccination
	Government and National Institute for Public Health recommend vaccination
	International organizations recommend vaccination
Media attention about the vaccine[6]	Traditional media positive (reference level)
	Traditional media negative
	Social and interactive media positive
	Social and interactive media negative
Out-of-pocket costs[7]	€0, €50, €100

Notes: [1]The scenario variables were the same for all alternatives in one choice set. [2]Defined as the proportion of population affected by the emerging disease, i.e. having symptoms. [3]Defined as the proportion of the infected population that had severe symptoms or outcomes (death, life-threatening events, hospitalisation and severe or permanent disability). [4]The attributes safety of the vaccine, advice about the vaccine and media attention about the vaccine were included in the latent class analysis as categorical variables. [5]Safety of the vaccine with regard to long-term severe side effects (death, life-threatening events, hospitalisation, severe or permanent disability, or side effects leading to birth defects in an unborn fetus). [6]Traditional media were defined as radio, newspapers and television. Social and interactive media were defined as blogs, Twitter and social network websites. [7]The levels presented in the Table are the selected levels for the Netherlands. Levels for the out-of-pocket costs attribute were converted to local currency of the other three countries and adapted according to the Organisation for Economic Co-operation and Development (OECD) price levels of May 2013 [26]. Levels of: 0 zloty, 120 zlotys, 240 zlotys for Poland; 0 euro, 45 euros and 90 euros for Spain and 0 kronor, 500 kronor, 1,000 kronor for Sweden.

The survey was first developed in Dutch and subsequently tested using think-a-loud interviews (n = 5) and a pen-and-paper pilot (n = 29). This resulted in some minor changes to the layout and phrasing of the Dutch survey. To be able to use the survey in the other countries, some further changes to the survey were made. For example, we adapted country naming, and currencies for the cost attribute based on Organisation for Economic Cooperation and Development (OECD) comparative price levels [25] of May 2013 [26]. Hereafter, the survey was translated into Polish, Spanish and Swedish. A second translator reviewed each translated survey. To minimise differences between the original Dutch and the translated versions of the survey and to check for inconsistencies, native speakers (speaking Dutch and the respective languages) translated each survey back into Dutch. In Spain, Sweden and

Poland, we asked 30 respondents per country to complete the adapted and back-translated survey online and to give their suggestions for improvement. No suggestions were given. More details of the DCE for the current study have been described elsewhere [27].

Data collection

An ISO certified market research company (ISO 26362 [28], ISO 20252 [29], and ISO 14001 [30]), was hired to administer the online survey. This company used their own panel to collect data in the Netherlands, while another company's panels were used to collect data in the other three countries. Both companies follow international quality standards for market research [31]. Panel members were emailed an URL to the survey. Quota sampling was used to ensure that samples were representative for each country based on age, sex, educational level and region. We aimed to have 500 completed surveys per country in order to obtain reliable outcomes [32]. All respondents gave informed consent before participating in the study and received a small financial incentive in local currency for their contribution to the study from the research company. The amount differed per country according to what is customary in the given country (e.g. Dutch respondents were paid 2.20 euros). Data collection took place between June and September 2013. A declaration of no objection was received from the Medical Ethics Committee of the Erasmus MC, University Medical Center Rotterdam (MEC-2012-263) after they reviewed the study protocol. According to Dutch legislation, the methodology of this study, a survey among volunteers of Internet panels, does not fall within the scope of the Medical Research Involving Human Subjects Act [33]. Although the aim of the study is of medical nature, respondents are not being subjected to any treatment or behavioural adjustments.

Data analysis

The choice observations resulting from the DCE were used to estimate the impact of pandemic scenario variables and vaccination programme attributes (independent variables) on the respondents' choices for vaccination or opting-out (dependent variable). A significant independent variable in this choice model indicates that the attribute or attribute level has a significant impact on vaccination preferences and the sign of the coefficient reflects whether this impact has a positive or negative effect. Note that pandemic scenario variables could only be included as an interaction effect, as the scenario was the same in the three alternatives presented in each choice set. Several types of discrete choice models can be estimated. We chose a latent class model, since this is a closed form model (i.e. does not rely on complex simulations) that can take the panel nature of the data into account (i.e. dependencies between choice observations by a single respondent) [34].

A latent class analysis assumes the existence of subgroups (i.e. classes) of respondents with homogenous preferences. The researcher pre-specifies the number of classes based on the best model fit using the Akaike Information Criterion (AIC) and sound interpretation of

classes. Class membership is latent in that the researcher does not determine who belongs to which class a priori. Instead, class membership is expressed by class probabilities that may depend on the respondent's characteristics. In addition to the choice model, we fitted a class membership model to test whether class membership is dependent on country of residence. Using the output of the class membership model, the class probabilities adjusted for country of residence can be calculated.

Calculation of the relative importance of the attributes enables a direct comparison of preferences between classes. The percentages represent the proportion of someone's preference (utility) that is based on that attribute. The relative importance can be calculated by dividing the difference in coefficient values between the highest and lowest level for a single attribute by the sum of the differences of all attributes for that class, considering interaction effects [35]. The mean expected uptake of a vaccine per class was calculated by taking the exponent of the total utility for vaccination divided by the exponent of utility of both vaccination and no vaccination. We were able to calculate these uptakes per country, by weighing the class-specific uptake with the class probabilities per country. The relative importance of the attributes and the expected vaccination uptake were calculated for two pandemic scenarios: a mild scenario in which 5% of the population gets the disease (susceptibility to the disease), and 5% of the sick people developing severe symptoms (severity of the disease), and a severe scenario in which 20% of the population gets the disease, and 75% of the sick people develops severe symptoms.

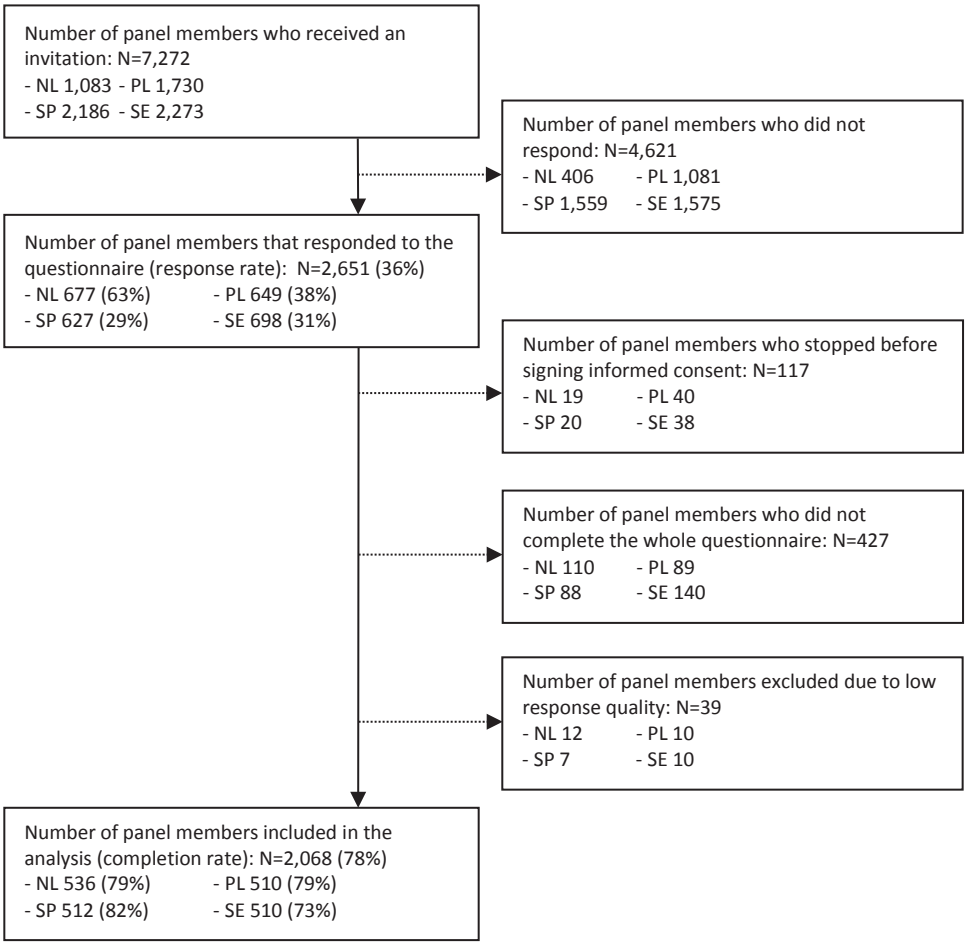
We used NLogit 4.0 software to estimate the latent class model and SPSS 21.0 software for all other analyses, such as chi-squared tests to compare proportions between countries.

RESULTS

Study population

In total 7,272 panel members were invited to participate in the study. Of these, 2,651 started the survey (response rates ranged from 29% (627/2,186) for Spanish panel members up to 63% (677/1,083) for Dutch panel members; Figure 1). Of those who started, 2,068 completed the survey, ranging from 73% (510/698) of Swedish panel members up to 82% (512/627) of Spanish panel members. The country samples were approximately representative regarding age, sex, educational level and region (Table 3). However, compared with national census data, lower educated Poles were slightly underrepresented as well as respondents from the western region of Spain.

Figure 1 | Response to the survey to investigate public preferences for vaccination programmes during pandemics caused by pathogens transmitted through respiratory droplets, Netherlands, Poland, Spain, and Sweden, 2013.



Abbreviations: *NL* Netherlands; *PL* Poland; *SE* Sweden; *SP* Spain.
Notes: [1]Low response quality was defined as completing the survey in less than 4 min.

Table 3 | Characteristics of respondents who completed the survey per country, Netherlands, Poland, Spain, and Sweden, 2013 (n = 2,068).

Characteristics	Netherlands (n = 536)			Poland (n = 510)			Spain (n = 512)			Sweden (n = 510)		
Age median (IQ range)	50 (35–64)			41 (28–55)			45 (31–57)			50 (35–59)		
	N	***%***	***%[1]***	***N***	***%***	***%[1]***	***N***	***%***	***%[1]***	***N***	***%***	***%[1]***
Age groups (years)												
18–24	49	9.1	*11*	95	19	*14*	59	12	*10*	58	11	*11*
25–34	78	15	*16*	95	19	*19*	95	19	*21*	69	14	*16*
35–44	84	16	*19*	101	20	*16*	97	19	*20*	77	15	*18*
45–54	107	20	*19*	90	18	*20*	79	15	*16*	112	22	*16*
≥55	218	41	*35*	129	25	*30*	182	36	*33*	194	38	*39*
Sex *(male)*	289	54	*49*	261	51	*48*	251	49	*49*	245	48	*49*
Country of birth is the country of interest	517	96	*NA*	502	98	*NA*	466	91	*NA*	440	86	*NA*
Educational level[2]												
Lower education	184	34	*34*	224	44	*52*	117	23	*23*	167	33	*33*
Average education	192	36	*40*	199	39	*34*	156	30	*31*	179	35	*34*
Higher education	160	30	*26*	87	17	*14*	239	47	*46*	164	32	*33*
Income[3]												
Low income	106	20	*NA*	133	26	*NA*	93	18	*NA*	120	24	*NA*
Average income	127	24	*NA*	127	25	*NA*	239	47	*NA*	256	50	*NA*
High income	181	34	*NA*	250	49	*NA*	180	35	*NA*	134	26	*NA*
Do not know or do not want to say	122	23	*NA*	0	0	*NA*	0	0	*NA*	0	0	*NA*
Religious *(yes)*	244	46	*NA*	403	79	*NA*	250	49	*NA*	191	37	*NA*
Working in healthcare *(yes)*	56	10	*NA*	20	4	*NA*	33	6	*NA*	48	9	*NA*
Perception of own health												
Worse health than average	41	8	*NA*	40	8	*NA*	36	7	*NA*	44	9	*NA*
Medium health	195	36	*NA*	165	32	*NA*	214	42	*NA*	151	30	*NA*
Better health than average	300	56	*NA*	305	60	*NA*	262	51	*NA*	315	62	*NA*

Seasonal influenza vaccine target group												
Yes	239	45	*NA*	85	17	*NA*	168	33	*NA*	136	27	*NA*
No	270	50	*NA*	382	75	*NA*	300	59	*NA*	321	63	*NA*
No, but receives vaccination via work	27	5	*NA*	43	8	*NA*	44	9	*NA*	53	10	*NA*
Received seasonal influenza vaccination last year *(yes, for persons belonging to target group)*	156	65	*NA*	34	40	*NA*	97	58	*NA*	56	41	*NA*

Abbreviations: *IQ* interquartile; *N*A not applicable.

Notes: [1]Census data per country. [2]Higher education was defined as: college, university, graduate degree; average education as: completed high school; and lower education as: all else, such as only elementary school or vocational education. [3]Income was defined as: low (< 23,000 euros), average (23,000–34,000 euros), high (> 34,000 euros) per year for the Dutch sample; low (< 2,000 zlotys), average (2,000–3,000 zlotys), high (> 3,000 zlotys) per month for the Polish sample; low (< 999 euros), average (1,000–2,000 euros),high (> 2,000 euros) per month for the Spanish sample; and low (< 175,000 kronor), medium (175,000–500,000 kronor), high (> 500,000 kronor) per year for the Swedish sample.

Table 4 | Regression coefficients for three latent classes based on responses to a survey investigating public preferences for vaccination programmes during pandemics caused by pathogens transmitted through respiratory droplets, Netherlands, Poland, Spain, and Sweden, 2013 (n = 2,068)[1,2,3].

Choice model	Class 1 Coefficient (p-value)	Class 1 SE	Class 2 Coefficient (p-value)	Class 2 SE	Class 3 Coefficient (p-value)	Class 3 SE
Constant (vaccination)	0.70 (***)	0.04	-0.79 (***)	0.03	-5.02 (***)	0.27
Effectiveness of vaccination (per 10%)	0.18 (***)	0.01	-0.03 (***)	0.01	0.06 (NS)	0.05
Side effects unknown, but expected to be safe (reference)	0.16 (Ref)	0.01	0.17 (Ref)	0.01	0.22 (Ref)	0.08
Side effects unknown, no experience yet	-0.16 (***)	0.01	-0.17 (***)	0.01	-0.22 (***)	0.08
Family and/or friends recommend (reference)[4]	-0.22 (Ref)	0.02	-0.14 (Ref)	0.02	0.33 (Ref)	0.16
Family and/or friends discourage	-0.34 (***)	0.02	-0.46 (***)	0.03	-0.41 (**)	0.19
Your doctor recommends	0.18(***)	0.02	0.40 (***)	0.02	0.50 (***)	0.15
Your doctor discourages	-0.47 (***)	0.02	-0.75 (***)	0.03	-1.05 (***)	0.28
Government & Public Health Institutions recommend	0.44 (***)	0.02	0.52 (***)	0.02	0.35 (**)	0.17
International organizations recommend	0.40 (***)	0.02	0.42 (***)	0.02	0.27 (*)	0.15
Traditional media is positive (reference)	0.03 (Ref)	0.01	0.22 (Ref)	0.02	0.33 (Ref)	0.12
Traditional media is negative	-0.12 (***)	0.02	-0.22 (***)	0.00	-0.41 (***)	0.15
Social / interactive media is positive	0.12 (***)	0.02	0.18 (***)	0.00	0.22 (*)	0.12
Social / interactive media is negative	-0.02 (NS)	0.02	-0.18 (***)	0.00	-0.14 (NS)	0.14
Out-of-pocket costs of the vaccine (per 10 euros)	-0.04 (***)	0.00	-0.13 (***)	0.00	-0.14 (***)	0.02
Interactions between attributes						
Effectiveness of vaccine (per 10%) x susceptibility to the disease (per 100 out of 1000 persons)	0.07 (***)	0.01	0.12 (***)	0.00	0.12 (***)	0.02
Effectiveness of vaccine (per 10%) x severity of the disease (per 10%)	0.01 (***)	0.00	0.02 (***)	0.00	0.01 (**)	0.00
Class membership model[5]						
Constant	-0.08 (NS)	0.10	0.00	0.00	-0.83 (***)	0.13
Netherlands (reference)	0.00 (Ref)	0.00	0.00	0.00	0.00 (Ref)	0.00
Poland	0.64 (***)	0.15	0.00	0.00	0.07 (NS)	0.20
Spain	0.60 (***)	0.15	0.00	0.00	0.12 (NS)	0.19
Sweden	-0.09 (NS)	0.16	0.00	0.00	0.86 (***)	0.17

Class probability[6]	**Proportion (RR)**	**Proportion (RR)**	**Proportion (RR)**
Average	0.44 (1.00)	0.35 (1.00)	0.21 (1.00)
Respondents from the Netherlands	0.39 (0.89)	0.42 (1.21)	0.18 (0.86)
Respondents from Poland	0.55 (1.24)	0.31 (0.89)	0.14 (0.69)
Respondents from Spain	0.53 (1.20)	0.32 (0.90)	0.16 (0.74)
Respondents from Sweden	0.30 (0.67)	0.35 (0.99)	0.36 (1.70)
Model fit[7,8]			
Akaike Information Criterion	1.54		
Pseudo-R2	0.30		

Abbreviations: *SE* standard error; *NA* not applicable; *NS* non-significant coefficient; *Ref* reference; *RR* relative risk.

Notes: [1]Effects coded variables used for the safety of the vaccine, advice about the vaccine, media attention about the vaccine. [2]The values of the vaccination programme attributes' reference levels equals the negative sum of the coefficients of the included attribute. [3] ***, **, * denotes significance at the 1% and 5% and 10% level respectively. [4]Note that for class 2 and 3, the recommendation of family and/or friends had a negative effect on utility. However, the utility is still positive compared with discouraging of family and/or friends. [5]Class 2 does not have parameters in the class membership model as the parameters of class 1 to 3 are relative to class 2. [6]The relative risks represent the relative probability of someone belonging to that class compared with the average class probability. [7]Note that the pseudo-R^2 is not the same as the R^2 that is used in a linear regression model. A pseudo-R^2 of 0.3–0.4 is equivalent to a R^2 between 0.6 and 0.8 [21]. [8]A model with 3 classes is presented in the Table. This model had significantly better fit compared with a model with 2 classes (AIC: 1.64, pseudo-R^2: 0.26). Although a latent class model with 4 classes had an improved fit (AIC: 1.50, pseudo-R^2: 0.32), we opted for a model with 3 classes to be able to explain the results to policymakers in a clear manner.

Respondents took a mean of 19 min (standard deviation: 31 min) to complete the survey. The majority of the respondents indicated that the survey topic was interesting or very interesting (81%; 1,677/2,068), and clear or very clear (74%; 1,528/2,068). A minority of respondents (9%; 179/2,068) found the survey hard or very hard to complete (ranging from 5% (28/510) for Poland to 13% (72/536) for the Netherlands). The proportion of choice sets in which the 'no vaccination' alternative was chosen was highest in the Swedish sample (51%; 4,145/(16*510=8,160)). The proportion of respondents that chose the 'no vaccination' alternative in all 16 choice sets was also higher in the Swedish sample (27% (136/510), $p < 0.01$) than elsewhere (10% for Poland (52/510) and Spain (54/512), and 11% (61/536) for the Netherlands). Additionally, the proportion of respondents that always opted for vaccination was lowest in the Swedish sample (16%; 81/510), and highest in the Spanish sample (31%; 161/512).

Latent class analysis

Three latent classes, numbered from one to three, were identified (Table 4). The average class probability was 0.44, 0.35 and 0.21, for class 1, 2, and 3 respectively. The country of residence partly explains class membership, which is an indication for preference heterogeneity between countries. Respondents from Poland and Spain had a significantly higher chance to belong to class 1 (0.55 and 0.53 respectively, $p < 0.01$) than respondents from other countries, those from the Netherlands had a significantly higher chance to belong to class 2 (0.42, $p < 0.01$), and those from Sweden to class 3 (0.36, $p < 0.01$).

Irrespective of the class they belonged to, respondents preferred a more effective vaccine that is expected to be safe, recommended by others, discussed positively in the media and with lower out-of-pocket costs, as can be seen by the positive and negative signs of the coefficients. The significant constant in all three classes indicates that, without considering any vaccination programme attributes, respondents of class 2 and 3 had a rather negative attitude towards vaccination, while respondents belonging to class 1 did not. Almost all vaccination programme attributes were significant. The positive recommendation of international organisations did not significantly explain preferences of respondents within class 3. The coefficient for social/interactive media attention was not significantly different from positive traditional media attention for respondents of class 3 (both positive and negative social/interactive media attention) and class 1 (only negative social/interactive media attention), meaning that social media only marginally influences respondents' preferences for vaccination. Significant interaction effects between both susceptibility to and severity of the disease, and effectiveness of the vaccine in all classes indicate that the preference for the level of effectiveness of a vaccine is dependent on the seriousness of the pandemic. In other words, the more serious the pandemic, while the effectiveness of a vaccination remains the same, the more the preference for vaccination increases relative to no vaccination.

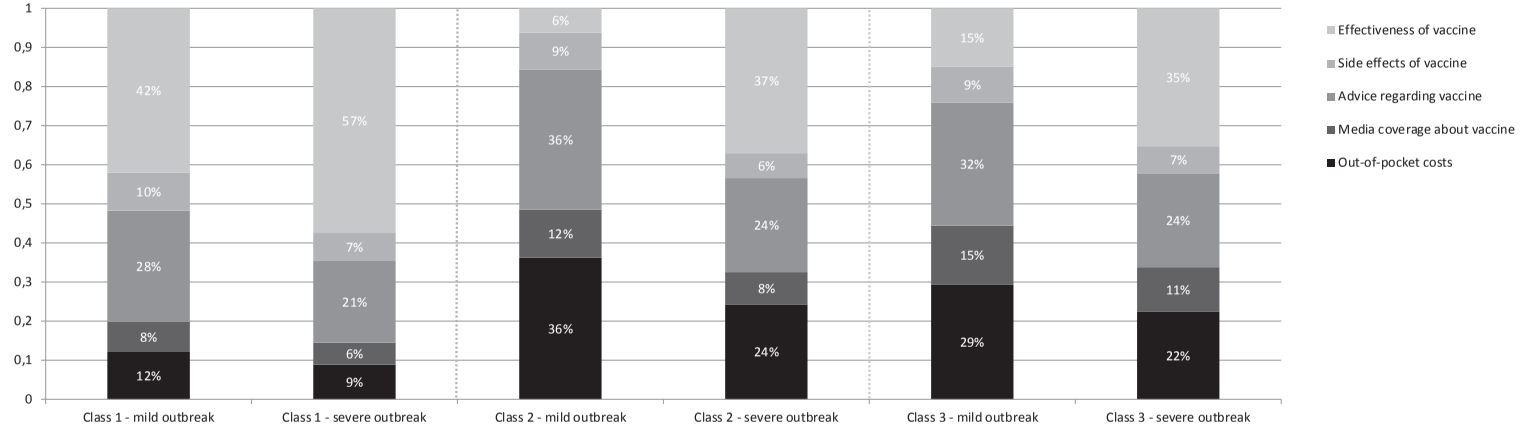
Relative importance

In the case of a mild scenario, the two most important attributes for class 2 and 3 were advice regarding vaccination and out-of-pocket costs, while effectiveness of the vaccine and advice regarding vaccination were the most important attributes for class 1 (Figure 2). Although advice regarding vaccination was important irrespective of class membership, for respondents belonging to class 3, the advice of friends and/or family and the advice of physicians were most important for vaccination choice (based on differences between coefficients of advice regarding vaccine), while the advice of both national and international health authorities was important for respondents belonging to class 1. Additionally, all respondents were more sensitive to advice against compared with advice in favour of vaccination. The relative importance of attributes varied with the seriousness of the pandemic scenario. Effectiveness was the most important attribute in the case of a severe scenario in all the latent classes and not only for respondents from class 1.

Predicted vaccine uptake

Assuming a realistic vaccination programme (i.e. a vaccination that is 70% effective, expected to be safe, recommended by family and/or friends, positively discussed in traditional media, and without out-of-pocket costs), the mean expected uptake in the case of a mild scenario was lowest for Swedish respondents with 43% (220/510; 95% confidence interval (CI): 40–47%)), followed by 54% (292/536; 95% CI: 51–58%) for Dutch respondents, 62% (318/512; 95% CI: 59–65%) for Spanish respondents, and highest for respondents from Poland with 63% (323/510, 95% CI: 60–66%). In the case of a mild scenario, advice regarding the vaccine and out-of-pocket costs had a relatively large impact on vaccination uptake in all countries, while media attention had little effect on uptake. For example, when out-of-pocket costs increased from 0 to 100 euros, the uptake decreased to 32% (163/510; 95% CI: 29–35%) for Swedish respondents, followed by 41% (222/536; 95% CI: 38–45%) for Dutch respondents, 51% (263/512; 95% CI: 48–55%) for Spanish respondents, and 53% (269/510; 95% CI: 49–56%) for Polish respondents. The uptake rates were expected to increase dramatically in the case of a severe scenario with up to 65% (331/510; 95% CI: 61–69%) for respondents from Sweden, and 82% (419/510; 95% CI: 80–85%) for respondents from Poland.

Figure 2 | Relative importance of vaccination programme attributes for respondents' decision to get vaccinated in the case of mild and severe pandemic scenarios caused by pathogens transmitted through respiratory droplets, Netherlands, Poland, Spain, and Sweden, 2013 (n = 2,068)[1,2,3].



Notes: [1]The percentages represent the proportion of someone's preference that is based on that attribute (utility). [2]A mild pandemic was defined as a pandemic in which 5% of the population gets the disease (pandemic scenario variable susceptibility), and 5% of the sick people developing severe symptoms (pandemic scenario variable severity). [3]A severe pandemic was defined as a pandemic in which 20% of the population gets the disease (pandemic scenario variable susceptibility), and 75% of the sick people develop severe symptoms scenario variable (pandemic scenario variable severity).

DISCUSSION

Statement of principal findings

In the case of a severe pandemic scenario, vaccine effectiveness was the most important characteristic determining vaccination preference in all countries. The body that advises a vaccine was found to strongly affect preferences in all countries as well, with respondents being more sensitive to advice against compared with advice in favour of vaccination. Preference heterogeneity between countries was substantial, especially in the case of a mild pandemic scenario; a strong effect on vaccine preferences was found for the advice of family and/or friends and the advice of physicians in Sweden, in contrast to Poland and Spain, where the advice of (international) health authorities was more important. Besides the vaccination advice, out-of-pocket costs were important for Dutch and Swedish respondents, while for respondents from Poland and Spain the effectiveness of the vaccine was important in case of a mild pandemic scenario. Irrespective of pandemic scenario or programme attributes, the predicted vaccination uptakes were lowest in Sweden, and highest in Poland.

Strengths and weaknesses of the study

So far, only a limited number of healthcare-related DCEs have quantitatively compared preferences between respondents from different countries and this is, to our best knowledge, done for the first time in the field of infectious diseases. An additional strength is the advanced analysis technique we used in this study. While already used extensively in the field of transport economics, latent class analysis has been used for only 3% of all health-related DCE analyses conducted between 2009 and 2012 [23]. A possible weakness of our study is that the preferences are stated and based on hypothetical pandemic scenarios. Respondents might have given socially desirable responses. It is not known to what extent the stated preferences differ from preferences during an actual pandemic. However, the external validity of the DCE method has been studied in other health related contexts, and results are encouraging with respect to prediction of preferences on an aggregate level [36,37]. In addition, the hypothetical nature of the study enabled us to compare preferences between different possible future pandemic scenarios. The findings might thus help to prepare for a future pandemic. Additionally, all coefficients had the expected sign, which suggests theoretical validity of the DCE [38]. Another possible weakness is the complexity of the choice sets, due to inclusion of risks as attributes. However, we thoroughly pilot tested the survey and, during the online survey, only a minority of respondents stated that they experienced problems completing the choice sets.

Results in relation to other studies

Our study showed that the expected vaccination uptake is largely dependent on the seriousness of a pandemic. This was also shown in previous studies, including studies conducted in the Netherlands, Poland, Spain and Sweden [39-45]. During the influenza

A(H1N1)pdm09 pandemic, the perceived vulnerability was low and respondents believed that they were less likely to become infected than other people [41,46]. This might have been one of the reasons for the lower than expected uptake during that pandemic with overall, 30%, 27% and 59% of the Dutch, Spanish and Swedish population respectively, having been vaccinated (Table 1). Interestingly, we found that Swedish respondents were least willing to get vaccinated in future influenza pandemics, both in mild and severe scenarios. As previous experiences are likely to influence future vaccination uptake [45], the difference between our study results and actual influenza A(H1N1)pdm09 vaccination coverage might be assigned to the negative experiences Swedish citizens had with vaccination during the 2009 pandemic. In Sweden, the controversy on the association between pandemic vaccines and narcolepsy is still ongoing [47]. In addition, Swedish respondents in the current study less often had received seasonal influenza vaccination in the previous year compared with e.g. Dutch respondents (41% vs 65%, Table 3). Research, conducted in the Netherlands, has shown that trust in health authorities is related to pandemic influenza vaccination uptake [48] and that it is necessary to build up and sustain trust before, during and after an influenza pandemic [16]. Furthermore, during the influenza A(H1N1)pdm09 pandemic Dutch and Swedish participants had more trust in healthcare professionals compared with Polish and Spanish participants [18]. Our research shows the same inter-country differences. Poland did not implement a national vaccination programme during the influenza A(H1N1) pdm09 pandemic [15,44] (Table 1). Seasonal influenza vaccination coverage is reported to be less than 10% for the target population older than 55 years [49]. Reported reasons for the Polish public to reject influenza (both seasonal and pandemic) vaccination include the low level of confidence in the quality and effectiveness of the vaccine [18,50]. Our finding that effectiveness of a pandemic vaccine had by far the strongest effect on vaccination choice of Polish respondents, confirmed this. The lowest seasonal influenza vaccination coverage contrasts with our finding that Polish respondents were more willing to get vaccinated than respondents from other countries. However, in our study, the level of effectiveness of the vaccine was presented to respondents as a known rate, which might explain why we estimated a higher vaccination uptake. Safety of the pandemic vaccine was not as dominant in the current study as in other studies [39,40]. The choice of attribute levels for our DCE might explain this difference in relative importance. We included realistic attribute levels, instead of presenting a certain vaccination risk (e.g. 1 in 100,000) to respondents. We also analysed safety as an interaction with the pandemic scenario variable 'severity of the disease', but with no meaningful outcome. We found almost no effect of social media attention (compared to traditional media) on pandemic vaccination preferences and predicted uptake. The objective framing of this attribute in the DCE survey might explain the finding. However, social media will likely be influential in future pandemics in other ways, e.g. by creating online applications that provide credible health information [51].

Implications for clinicians and policymakers

Our results show that seriousness of a pandemic influences vaccination uptake dramatically. In order to increase pandemic vaccination coverage, it is essential that susceptible people feel susceptible and perceive the pandemic as a serious threat. This can be achieved, for example, by honest and open communication regarding the seriousness of the pandemic, and avoiding conflicting messages and information overload [17,52] and by providing public health messages that include descriptive and injunctive normative information [53,54]. The WHO Regional Office for Europe and the European Centre for Disease Prevention and Control (ECDC) recommend more flexible pandemic preparedness planning, i.e. planning that takes into account different pandemic scenarios [8,9,19]. Findings of our study may facilitate responses to future influenza pandemics with different levels of severity, as our study provides the option to calculate the expected vaccination uptake for different pandemic scenarios, and provides insights into how several vaccination programme attributes influence these uptakes. Additionally, our study also shows that the availability of an effective pandemic vaccine is of paramount importance in order to reach certain coverage levels. Unfortunately, such a highly effective vaccine might not be available due to the crisis situation that is inherent to a pandemic, or proof that the vaccine is effective might be lacking as time is usually limited. In addition, due to contracts or limited availability of vaccines, there are usually only one or two different vaccines available for policymakers to choose from. For all countries, given the high impact of vaccine effectiveness on vaccination preferences, it is therefore important that there is open communication regarding the expected effectiveness, so that the public can make an informed choice whether to get vaccinated or not. The vaccination programme attributes that can be influenced by policymakers directly are out-of-pocket costs and how/what to communicate. As our results show that by whom a vaccine is advised had a different effect on uptake in the included countries, it is important that during future pandemics the responsible authorities align with other important stakeholders in the country and communicate in a coordinated manner.

Unanswered questions and further research

We found differences in preferences for pandemic vaccinations between difference European countries. Further research could focus on differences within these countries, e.g. whether preferences of those who previously received seasonal influenza vaccination differ from preferences of those who had not, as previous research shows that the uptake of seasonal influenza vaccination was positively associated with influenza A(H1N1)pdm09 vaccination decision-making [39,55,56]. Additionally, future research could focus on subgroups of the population, such as healthcare workers or under-vaccinated groups. It is unknown whether preferences differ between countries within the same geographical area of Europe. Therefore, it might be useful to conduct the same DCE in other European countries as well. Unfortunately, timely access to vaccinations is not self-evident [57]. It is not known in advance which respiratory pathogen will cause a next pandemic and

production capacities might be inadequate. In the case of an influenza pandemic, other preventive measures such as quarantine, and antiviral drugs might be helpful to limit the spread of the virus during the first phase [10]. Further research into preferences for other preventive measures, and differences herein across European countries, using the DCE methodology is thus recommended. Moreover, the DCE methodology could also be used to study motivations and barriers for vaccinations other than pandemic vaccination among different countries.

REFERENCES

1. Taubenberger JK, Morens DM. 1918 Influenza: the mother of all pandemics. Emerg Infect Dis. 2006;12:15-22.
2. The influenza A(H1N1) pandemic in Europe. Stockholm: European Centre for Disease Prevention and Control (ECDC); 2010. Available from: http://ecdc.europa.eu/en/publications/Publications/101108_SPR_pandemic_experience.pdf.
3. Kilbourne ED. Influenza pandemics of the 20th century. Emerg Infect Dis. 2006;12(1):9-14.
4. Nicoll A, Ammon A, Amato Gauci A, et al. Experience and lessons from surveillance and studies of the 2009 pandemic in Europe. Public Health. 2010;124:14-23.
5. Known facts and known unknowns. Stockholm: European Centre for Disease Prevention and Control (ECDC); 2010. Available from: http://ecdc.europa.eu/en/healthtopics/Documents/0905_Pandemic_Influenza_Known_Facts_and_Known_Unknowns.pdf. [Last accessed 21 10 2015].
6. WHO global influenza preparedness plan: the role of WHO and recommendations for national measures before and during pandemics. Geneva: World Health Organization (WHO); 2005. Available from: http://www.who.int/csr/resources/publications/influenza/WHO_CDS_CSR_GIP_2005_5.pdf.
7. Fineberg HV. Pandemic preparedness and response--lessons from the H1N1 influenza of 2009. N Engl J Med. 2014;370:1335-42.
8. Nicoll A, Brown C, Karcher F, et al. Developing pandemic preparedness in Europe in the 21st century: experience, evolution and next steps. Bull World Health Organ. 2012;90:311-7.
9. Key changes to pandemic plans by Member States of the WHO European Region based on lessons learnt from the 2009 pandemic. World Health Organization Regional Office for Europe (WHO/Europe), European Centre for Disease Prevention and Control (ECDC); 2012. Available from: http://www.euro.who.int/__data/assets/pdf_file/0006/161664/ECDC_WHO_EURO_PiP-Workshops-Summary-Report_FINAL-26032012.pdf?ua=1.
10. Longini IM Jr, Nizam A, Xu S, et al. Containing pandemic influenza at the source. Science. 2005;309:1083-7.
11. Ferguson NM, Cummings DA, Fraser C, et al. Strategies for mitigating an influenza pandemic. Nature. 2006;442:448-52.
12. Germann TC, Kadau K, Longini IM Jr, et al. Mitigation strategies for pandemic influenza in the United States. Proc Natl Acad Sci USA. 2006;103:5935-40.
13. Lugnér AK, van Boven M, de Vries R, et al. Cost effectiveness of vaccination against pandemic influenza in European countries: mathematical modelling analysis. BMJ. 2012;345:e4445.
14. Valenciano M, Kissling E, Cohen JM, et al. Estimates of pandemic influenza vaccine effectiveness in Europe, 2009-2010: results of Influenza Monitoring Vaccine Effectiveness in Europe (I-MOVE) multicentre case-control study. PLoS Med. 2011;8:e1000388.
15. Mereckiene J, Cotter S, Weber JT, et al. Influenza A(H1N1)pdm09 vaccination policies and coverage in Europe. Euro Surveill. 2012;17:20064.
16. Gesser-Edelsburg A, Mordini E, James JJ, et al. Risk Communication Recommendations and Implementation During Emerging Infectious Diseases: A Case Study of the 2009 H1N1 Influenza Pandemic. Disaster Med Public Health Prep. 2014;8: 158-69.
17. Larson HJ, Cooper LZ, Eskola J, et al. Addressing the vaccine confidence gap. Lancet. 2011;378:526-35.
18. Influenza H1N1 Analytical Report - Eurobarometer on Influenza H1N1. The Gallup Organization; 2010. Available from: http://ec.europa.eu/public_opinion/flash/fl_287_en.pdf.
19. O'Flanagan D, Cotter S, Mereckiene J. Pandemic A(H1N1) 2009 Influenza Vaccination Survey, Influenza season 2009/2010. VENICE II Consortium; 2011. Available from: http://venice.cineca.org/Final_Report_VENICE_Pandemic_Influenza_2009.pdf.
20. Ryan M. Discrete choice experiments in health care. BMJ. 2004;328:360-1.
21. Hensher DA, Rose JM, Greene WH. Applied choice analysis: a primer. New York: Cambridge University Press; 2005.

22. Luce D, Tukey J. Simultaneous conjoint measurement: a new type of fundamental measurement. J Math Psychol. 1964;1:1-27.
23. Clark MD, Determann D, Petrou S, et al. Discrete choice experiments in health economics: a review of the literature. Pharmacoeconomics. 2014;32:883-902.
24. Reed Johnson F, Lancsar E, Marshall D, et al. Constructing experimental designs for discrete-choice experiments: report of the ISPOR Conjoint Analysis Experimental Design Good Research Practices Task Force. Value Health. 2013;16:3-13.
25. Monthly comparative price levels. Organisation for Economic Co-operation and Development (OECD), 2015. Available from: http://stats.oecd.org/Index.aspx?QueryId=24057. [Last accessed 27 10 2015].
26. Main economic indicators. Organisation for Economic Co-operation and Development (OECD); 2013. Available from: http://dx.doi.org/10.1787/mei-v2013-5-en.
27. Determann D, Korfage IJ, Lambooij MS, et al. Acceptance of vaccinations in pandemic outbreaks: a discrete choice experiment. PLoS One. 2014;9:e102505.
28. ISO 26362:2009(en) Access panels in market, opinion and social research. International Organization for Standardization (ISO). Available from: www.iso.org/obp/ui/#iso:std:iso:26362:ed-1:v1:en. [Last accessed 29 09 2015].
29. ISO 20252:2012(en) Market, opinion, and social research. International Organization for Standardization (ISO). Available from: www.iso.org/obp/ui/#iso:std:iso:20252:ed-2:v1:en. [Last accessed 29 09 2015].
30. ISO 14001:2015(en) Environmental management systems. International Organization for Standardization (ISO). Available from: www.iso.org/obp/ui/#iso:std:iso:14001:ed-3:v1:en. [Last accessed 14 10 2015].
31. The ICC/ESOMAR code on Market and Social Research. Available from: www.esomar.org/knowledge-and-standards/codes-and-guidelines.php. [Last accessed 14 10 2015].
32. Orme BK. Getting started with Conjoint Analysis: Strategies for Product Design and Pricing Research. Madison, WI: Research Publishers LLC; 2006.
33. [Medical Research Involving Human Subjects Act]. Wet medisch-wetenschappelijk onderzoek met mensen. Available from: http://wetten.overheid.nl/BWBR0009408. [Last accessed 29 09 2015]. *In Dutch.*
34. Greene WH, Hensher DA. A latent class model for discrete choice analysis: contrasts with mixed logit. Transport Res B-Meth. 2003;37:681-98.
35. Louviere JH, Hensher DA, Swait JD. Stated choice models: analysis and application. Cambridge: Cambridge University Press; 2000.
36. Kesternich I, Heiss F, McFadden D, et al. Suit the action to the word, the word to the action: Hypothetical choices and real decisions in Medicare Part D. J Health Econ. 2013;32:1313-24.
37. Linley WG, Hughes DA. Decision-makers' preferences for approving new medicines in Wales: a discrete-choice experiment with assessment of external validity. Pharmacoeconomics. 2013;31:345-55.
38. Lancsar E, Louviere J. Conducting discrete choice experiments to inform healthcare decision making: a user's guide. Pharmacoeconomics. 2008;26:661-77.
39. Bish A, Yardley L, Nicoll A, et al. Factors associated with uptake of vaccination against pandemic influenza: a systematic review. Vaccine. 2011;29:6472-84.
40. Sadique MZ, Devlin N, Edmunds WJ, et al. The effect of perceived risks on the demand for vaccination: results from a discrete choice experiment. PLoS One. 2013;8:e54149.
41. Bults M, Beaujean DJ, de Zwart O, et al. [Mexican flu: risk perception in the general public, precautionary measures and trust in information provided by the government]. Mexicaanse griep: risicoperceptie bij de bevolking, eigen maatregelen en vertrouwen in overheidsinformatie. Ned Tijdschr Geneeskd. 2010;154:A1686. *In Dutch.*
42. Zijtregtop EA, Wilschut J, Koelma N, et al. Which factors are important in adults' uptake of a (pre) pandemic influenza vaccine? Vaccine. 2009;28:207-27.
43. Agüero F, Adell MN, Pérez Giménez A, et al. Adoption of preventive measures during and after the 2009 influenza A (H1N1) virus pandemic peak in Spain. Prev Med. 2011;53:203-6.

44. Brydak LB, Woźniak Kosek A, Nitsch-Osuch A. Influenza vaccines and vaccinations in Poland - past, present and future. Med Sci Monit. 2012;18:RA166-71.
45. Börjesson M, Enander A. Perceptions and sociodemographic factors influencing vaccination uptake and precautionary behaviours in response to the A/H1N1 influenza in Sweden. Scand J Public Health. 2014;42:215-22.
46. Seale H, Heywood AE, McLaws ML, et al. Why do I need it? I am not at risk! Public perceptions towards the pandemic (H1N1) 2009 vaccine. BMC Infect Dis. 2010;10:99.
47. Barker CI, Snape MD. Pandemic influenza A H1N1 vaccines and narcolepsy: vaccine safety surveillance in action. Lancet Infect Dis. 2014;14:227-38.
48. Van der Weerd W, Timmermans DR, Beaujean DJ, et al. Monitoring the level of government trust, risk perception and intention of the general public to adopt protective measures during the influenza A (H1N1) pandemic in The Netherlands. BMC Public Health. 2011;11:575.
49. Mereckiene J, Cotter S, Nicoll A, et al.; VENICE project gatekeepers group. Seasonal influenza immunisation in Europe. Overview of recommendations and vaccination coverage for three seasons: pre-pandemic (2008/09), pandemic (2009/10) and post-pandemic (2010/11). Euro Surveill. 2014;19:20780.
50. Kardas P, Zasowska A, Dec J, et al. Reasons for low influenza vaccination coverage: cross-sectional survey in Poland. Croat Med J. 2011;52:126-33.
51. Merchant RM, Elmer S, Lurie N. Integrating social media into emergency-preparedness efforts. N Engl J Med. 2011;365:289-91.
52. Lin L, Savoia E, Agboola F, et al. What have we learned about communication inequalities during the H1N1 pandemic: a systematic review of the literature. BMC Public Health. 2014;14:484.
53. Nolan JM, Schultz PW, Cialdini RB, et al. Normative social influence is underdetected. Pers Soc Psychol Bull. 2008;34:913-23.
54. Schultz PW, Nolan JM, Cialdini RB, et al. The constructive, destructive, and reconstructive power of social norms. Psychol Sci. 2007;18:429-34.
55. Brien S, Kwong JC, Buckeridge DL. The determinants of 2009 pandemic A/H1N1 influenza vaccination: a systematic review. Vaccine. 2012;30:1255-64.
56. Nguyen T, Henningsen KH, Brehaut JC, et al. Acceptance of a pandemic influenza vaccine: a systematic review of surveys of the general public. Infect Drug Resist. 2011;4:197-207.
57. Evolution of a pandemic: A(H1N1) 2009, April 2009 - August 2010. Geneva: World Health Organization (WHO); 2013. Available from: http://apps.who.int/iris/bitstream/10665/78414/1/9789241503051_eng.pdf?ua=1.
58. Mereckiene J, Cotter S, O'Flanagan D. National seasonal influenza vaccination survey for 2012-13 influenza season in EU/EEA (provisional data). Venice III Consortium; 2014. Available from: http://venice.cineca.org/VENICE_Seasonal_Influenza_2012-13_v10.pdf.

5

Acceptance of vaccinations in pandemic outbreaks: a discrete choice experiment

Domino Determann
Ida J. Korfage
Mattijs S. Lambooij
Michiel C. Bliemer
Jan Hendrik Richardus
Ewout W. Steyerberg
Esther W. de Bekker-Grob

PLoS ONE. 2014;9(7):e102505
doi: 10.1371/journal.pone.0102505

ABSTRACT

Background: Preventive measures are essential to limit the spread of new viruses; their uptake is key to their success. However, the vaccination uptake in pandemic outbreaks is often low. We aim to elicit how disease and vaccination characteristics determine preferences of the general public for new pandemic vaccinations.

Methods: In an internet-based discrete choice experiment (DCE) a representative sample of 536 participants (49% participation rate) from the Dutch population was asked for their preference for vaccination programs in hypothetical communicable disease outbreaks. We used scenarios based on two disease characteristics (susceptibility to and severity of the disease) and five vaccination program characteristics (effectiveness, safety, advice regarding vaccination, media attention, and out-of-pocket costs). The DCE design was based on a literature review, expert interviews and focus group discussions. A panel latent class logit model was used to estimate which trade-offs individuals were willing to make.

Results: All above mentioned characteristics proved to influence respondents' preferences for vaccination. Preference heterogeneity was substantial. Females who stated that they were never in favour of vaccination made different trade-offs than males who stated that they were (possibly) willing to get vaccinated. As expected, respondents preferred and were willing to pay more for more effective vaccines, especially if the outbreak was more serious (€6–€39 for a 10% more effective vaccine). Changes in effectiveness, out-of-pocket costs and in the body that advises the vaccine all substantially influenced the predicted uptake.

Conclusions: We conclude that various disease and vaccination program characteristics influence respondents' preferences for pandemic vaccination programs. Agencies responsible for preventive measures during pandemics can use the knowledge that out-of-pocket costs and the way advice is given affect vaccination uptake to improve their plans for future pandemic outbreaks. The preference heterogeneity shows that information regarding vaccination needs to be targeted differently depending on gender and willingness to get vaccinated.

INTRODUCTION

Worldwide viral infection outbreaks with, e.g. Influenza A(H1N1), SARS and H5N1 avian influenza, have been of serious impact in the past [1]. If a new outbreak would occur, the global spread is likely to be very rapid due to increased travel and urbanization [2]. Extrapolation of the 1918–1920 avian influenza pandemic mortality rates indicates that 62 million people would be killed if a similar pandemic would happen these days [3]. Preventive measures, such as social distancing measures or vaccination programs, are very important in limiting the spread of new viruses [4,5]. However, the lack of willingness to act according to such measures in crisis situations has proven to be a major issue in the European Union [6]. Consequently, it is important to have insights into what motivates individual people to decide for or against vaccination. If motivations are known, these can be addressed in pandemic preparedness plans and vaccination strategies to increase vaccination rates and thus reduce the spread of viral outbreaks. Furthermore, insight in motivations can lead to an accurate prediction of the uptake of vaccinations, which is helpful when implementing vaccination programs.

Various studies have been conducted to explore reasons why individual members of the general public accepted or declined pandemic vaccinations, especially focusing on the Influenza A(H1N1) pandemic of 2009 [6–8]. These studies showed that participation in vaccination programs is based on weighing the burden of the vaccination (e.g. risk of side effects), against its potential benefits (e.g. reduce the risk of infection), in a given context (e.g. severity of first cases of the disease). Several European countries reported that public perception factors, such as poor confidence in the need for the vaccine and concerns about the relatively new vaccine, may have contributed to the low vaccination coverage rates during the Influenza A(H1N1) pandemic of 2009 [9]. Despite the presence of studies investigating reasons of members of the general public to get vaccinated or not, quantitative studies that assess the relative importance of these reasons are lacking. It is precisely this information that is needed to make highly effective health care policy plans regarding pandemic outbreaks and vaccinations.

The aim of this study is to investigate the preferences of the general population for pandemic vaccinations quantitatively. Additionally, we aim to calculate the expected uptake of base case vaccination programs for certain hypothetical outbreaks. The current study is conducted within the scope of the project Effective Communication in Outbreak Management: development of an evidence-based tool for Europe (ECOM, http://www.ecomeu.info/). This project aims to develop an evidence-based behavioural and communication strategy for health professionals and agencies throughout Europe in case of major outbreaks, by integration of social, behavioural, communication, and media sciences.

METHODS

Ethics Statement

A declaration of no objection was received from the Medical Ethics Committee of the Erasmus MC, University Medical Center Rotterdam (MEC-2012-263) after they reviewed the study protocol. The methodology of this study, a survey amongst healthy volunteers of an internet panel, does not fall within the scope of the Medical Research Involving Human Subjects Act (in Dutch: WMO). Although the aim of the study is of medical nature, participants are not being subjected to any treatment or behavioural adjustments.

Discrete choice experiments

DCE methodology is a survey-based stated preference technique to quantitatively investigate individual preferences. DCEs have been widely used in health care to examine stakeholder preferences [10, 11] and have been previously used to examine preferences for non-emergency vaccination programs, such as Human Papilloma Virus (HPV) vaccinations and seasonal influenza vaccinations [12, 13]. In DCEs, it is assumed that a medical intervention, such as a vaccination program, can be described by its characteristics (attributes; e.g. effectiveness of a vaccine, safety of the vaccine, and costs of the vaccine). Those characteristics are further specified by variants of that characteristic (attribute levels; e.g. for effectiveness of a vaccine: 30%, 50%, 70% and 90% effective). A second assumption is that the individual's preference for a medical intervention is determined by the levels of those attributes [14]. The relative importance of attributes can be assessed by presenting respondents a series of questions in which they are asked to choose a preferred alternative from a set of two or more hypothetical intervention alternatives with varying combinations of attribute levels [15]. DCEs are based on Lancaster's consumer theory [16] and random utility theory (RUT) [17] which assume that an individual acts rationally and always chooses the alternative with the highest level of utility. We followed recent guidelines for good DCE practice [18, 19].

Selection of attributes and attribute levels

Only a limited number of attributes and attribute levels can be used in a DCE, since otherwise the precision and reliability of the results will decrease. On the other hand, one also needs to include all relevant attributes and attribute levels to avoid that respondents make significant inferences on omitted attributes or levels [19,20].

To obtain insights into possible attributes and their levels to be included in this DCE, we conducted a strategic literature search in three databases (searching for literature related to DCEs and/or vaccination preferences in PubMed, Embase and PsychINFO), semi-structured expert interviews and a focus group study. For the expert interviews, we have spoken to nine experts of different relevant fields, e.g. infectious diseases, vaccinations, preventive

behaviour and implementation of preventive measures. For the focus group study, we conducted seven focus group discussions with the general population from the Netherlands; targeting urban populations (two groups); populations of more rural areas (two groups); and ethnic minorities in the Netherlands (three groups). Eligible participants were recruited by a research company and via the network of a researcher of the department of Public Health of the Erasmus MC, University Medical Center Rotterdam, using purposive sampling to ensure a diverse sample. Participants were informed that they would receive a financial incentive (40 euros) for their contribution and to cover travel costs and they were informed that the data would be analysed anonymously. All participants gave written informed consent prior to the discussions. All focus groups were audio taped, transcribed verbatim and anonymously. The transcripts were analysed using thematic analysis using NVivo Software (version 10, http://www.qsrinternational.com). The focus group study approach was included in the study protocol for which a declaration of no objection was received from the Medical Ethics Committee of the Erasmus MC, University Medical Center Rotterdam. We used a topic list based on the literature search and on two theoretic models, i.e. the Health Belief Model [21] and to a lesser extent the Protection Motivation Theory [22], to structure the focus group discussions on outbreaks of new diseases and preventive measures. These models assume that people react to a perceived threat, by performing some action. The level of threat depends on the perceived susceptibility to a disease and the perceived severity of a disease. People weigh this threat to perceived benefits (such as effectiveness) and barriers (such as costs) of actions. The model assumes that also other factors influence someone's intention to take some action, such as cues to action (e.g. media attention) and variables (age, sex, peer pressure etc.). We used these models as a base for the topic list because of their largely empirically tested ability to explain and predict intention of and complying with preventive medical care recommendations, including vaccinations [7, 23, 24]. Additionally, during the focus group discussions, participants were asked to write down and rank the most important reasons for them to get vaccinated during future pandemic outbreaks.

Using these results and through extensive discussion with ECOM project members, we selected two disease specific scenario variables and five vaccination program attributes and their corresponding levels (Table 1). For each scenario (which is a combination of the susceptibility to the disease and severity of the disease), three alternatives were presented, namely (i) No vaccination, (ii) Vaccination A, and (iii) Vaccination B, where the latter two are represented by combinations of effectiveness, safety, advice, media and out-of-pocket costs. We aimed at selecting a sufficient wide range of attribute levels that are realistic now and will remain so in the near future and levels that were relevant to policy as well as plausible and understandable for the respondents. Furthermore, for each continuous attribute we selected at least three levels to be able to test for non-linear relationships.

Table 1 | Scenario variables, vaccination program attributes and their levels included in the DCE survey.

Scenario variables	Levels
Susceptibility to the disease[1]	5%
	10%
	20%
Severity of the disease[2]	5%
	25%
	50%
	75%
Vaccination program attributes	**Levels**
Effectiveness of vaccine	30%
	50%
	70%
	90%
Safety of the vaccine[3,4]	Unknown, expected to be safe (reference level)
	Unknown, no experience with similar vaccines yet
Advice regarding the vaccine[3]	Family and/or friends recommend vaccination (reference level)
	Family and/or friends discourage vaccination
	Your doctor recommends vaccination
	Your doctor discourages vaccination
	Dutch government & RIVM recommend vaccination
	International organizations recommend vaccination
Media coverage about the vaccine[3]	Traditional media[5] positive (reference level)
	Traditional media[5] negative
	Social / interactive media[6] positive
	Social / interactive media[6] negative
Out-of-pocket costs	€0
	€50
	€100

Notes: Levels of the no vaccination option were defined as: not applicable (n.a.), no side effects, n.a., n.a., €0 respectively. The scenario variables were the same across all alternatives in one choice set. Abbreviation used: RIVM = Dutch abbreviation of National Institute for Public Health and the Environment. [1]Defined as the proportion of population infected with new disease, i.e. having symptoms. [2]Defined as the proportion of infected population that suffered severe symptoms (death, life-threatening events, hospitalization and severe or permanent handicap). [3]The attributes 'safety of the vaccine', 'advice about the vaccine' and 'media coverage about the vaccine' entered the analysis as categorical variables. [4]Long term severe side effects (death, life-threatening events, hospitalization, severe or permanent handicap, or side effects leading to birth defects to an unborn foetus). Before the start of the choice tasks, respondents were informed that on the short term, vaccinations resulted in mild side effects only. [5]Traditional media were defined as: radio, newspapers and television. [6]Social / interactive media were defined as: blogs, Twitter and social network websites.

Study design and questionnaire

If all combinations of attribute levels were to be presented in choice sets, this would have led to 576 ($2^1 * 3^1 * 4^2 * 6^1$) hypothetical vaccination alternatives for 12 ($3^1 * 4^1$) different disease outbreaks (scenarios). As it is not feasible to present a single individual with all these

scenarios and alternatives (i.e. full factorial design), a subset of scenarios and alternatives (i.e. fractional factorial design) was generated [25]. Zero priors for all categorical variables and best-guess priors for all linear variables were used to generate an efficient design by maximizing D-efficiency (using Ngene software, version 1.1.1, http://www.choice-metrics.com/). With this design we were able to estimate all main effects and a number of two-way and higher order interactions between attributes. Presenting a single individual with a large amount of choice sets is expected to result in a lower response rate and/or lower response reliability [26]. To reduce the burden on respondents, a blocked design was used [15], which resulted in dividing the 48 choice sets of the efficient design into 3 questionnaire versions containing 16 choice sets each in which we ensured sufficient variation in attribute levels by finding blocks with near attribute level balance.

Each questionnaire started with the introduction of a hypothetical scenario (Additional file 1). To facilitate comprehension of the DCE task, respondents were provided with detailed information about the attributes and attribute levels as well as with a clearly explained example of a choice task prior to preference elicitation. The main part of each questionnaire comprised 16 choice sets. In each choice set, respondents first received some additional information about the disease (i.e., the two scenario variables). Choice sets consisted of two unlabelled vaccination alternatives (vaccination A and vaccination B) and one opt-out alterative (see additional file 2 for a screenshot of a choice set). This opt-out was necessary since, as in real life, respondents are not obliged to take a vaccination. Respondents were asked to consider all three alternatives in a choice set as realistic alternatives and to choose the option that appealed most to them in the given situation.

Attributes needed to be described as clearly as possible in the choice sets since previous research has shown that respondents may have difficulties with interpreting probabilities [27] and that framing effects can influence DCE results [28–30]. Therefore, we included graphs to demonstrate percentages and rates, used realistic presentation of attributes (e.g. integers when discussing rates that included humans), and used cost as the last attribute. Furthermore, experts in the field of risk communication advised us on how to present the choice sets in this DCE. For example, we were advised to use the same type of graphs to present risks across both scenario variables and attributes. The last section of each questionnaire included questions on socio-demographic data and questions on previous experiences with vaccination. It also contained questions assessing experienced difficulty of the questionnaire (five-point scale). The questionnaire was presented to respondents in Dutch.

In order to test the survey, we conducted a formal pen and paper pilot with 29 respondents in the Netherlands. Additionally, we conducted five think-a-loud interviews [31] to qualitatively test for any problems in interpretation, for the understanding of the questions

and to indicate whether respondents were providing a meaningful response. This resulted in minor changes to the layout and phrasing of the questionnaire. There was no need to adapt the selected combinations of scenarios or attribute levels of the DCE design. Since there were some adaptations to the questionnaire, data of these pre-tests were not included in the final analysis. The questionnaire is available from the authors on request.

To check the convergent validity of the DCE, we asked respondents to rank the five vaccination program attributes from most important to least important. External validation was not possible since we were using a hypothetical disease outbreak.

Data collection

A market research company (Flycatcher) was hired to administer the online questionnaire to a representative sample of the general adult population of the Netherlands. Their online panel comprises 16,000 members and is ISO certified (ISO-26363). Recruitment of potential new members is done by digital media, paper invitations, face-to-face meetings and via intermediates. Assuming a participation rate of 50%, a random sub sample of 1,083 adult panel members (see sample size calculation below) was emailed a link to the questionnaire to participate in the current study. Quota sampling was used to ensure even distributions with respect to age, gender, education and region. A further quota was applied to each 'questionnaire version' to ensure comparable numbers of respondents in each of the three blocks of the design. Progress bars and error messages were incorporated to encourage completion. After completing the questionnaire, respondents were given the opportunity to comment on the questionnaire or topic at hand by filling out the free text question. The questionnaire was online for twelve days in June 2013, when the target number of 500 respondents was reached. All panel members gave informed consent prior to participating in the study and received a small incentive (€2.20, in the form of credits) for completion of the questionnaire.

Sample size calculation

The mean sample size for DCE studies in health care published between 2005 and 2008 was 259, with nearly 40% of the sample sizes in the range of 100 to 300 respondents [32]. No adequate statistical methods exist to determine sample sizes for DCEs. Therefore, the rules of thumb as suggested by Orme [33] are frequently used. These rules recommend sample sizes for DCEs to be at least 300 respondents and suggest that also the number of tasks and alternatives should be taken into account when determining sample sizes. Based on this information, we aimed to have at least 500 respondents completing the questionnaire.

Statistical analysis

To assess preference heterogeneity, we used a latent class model to analyse the DCE data. A latent class model [34,35] can be used to identify the existence and the number of

segments or classes in the population (i.e. identifying different utility (preference) functions across unobserved subgroups). Class membership is latent (i.e., unobserved) because each respondent belongs to each class up to a modelled probability and not deterministically assigned by the analyst a priori. The model is flexible in that the probability that sampled respondents belong to a particular class can be linked to covariates (e.g. age, gender), hence allowing for some understanding as to the make-up of the various class segments [34].

To account for the panel nature of the data, with each respondent completing 16 choice tasks, we used a panel version of the latent class model. In order to determine the number of classes, we selected the model with the best fit. We tested a number of different specifications for the utility function (e.g., categorical or numerical attribute levels, linearity, two-way interactions between all attributes and several attribute transformations, see Additional file 3 for specifications of the functions) and selected the model with the lowest Akaike Information Criterion (AIC).

The latent class model estimates parameters in a class assignment model (which includes socio-demographic variables and thereby expresses the likelihood of a certain individual belonging to a certain class) and class-specific coefficients for each attribute (or interaction of attributes and scenario variables) in the utility function. For the class-specific coefficients and interactions, the statistical significance of a coefficient (P-value ≤ 0.05) indicated that, conditional on belonging to that class, respondents considered the attribute important when making stated choices. In terms of the class assignment parameters, statistically significant parameter estimates indicate that the covariate can be used to distinguish between the different classes. For example, if the covariate male gender is negatively and significantly associated with a particular class in the assignment model, then this is indicative that men are less likely to belong to that particular class than women.

The sign of the coefficient reflects whether the attribute had a positive or negative effect on utility. The value of each coefficient represents the importance respondents assign to an attribute (level). However, different attributes utilize different units of measurement. For example the coefficient ‘effectiveness of the vaccine’ represented the importance per 10% protection rate. When looking at a vaccine that generates a 90% protection rate, the coefficient needs to be multiplied 9 times (9 times coefficient of ‘effectiveness of the vaccine’ of 10% = coefficient of ‘effectiveness of the vaccine’ of 90%).

We calculated class specific importance scores (IS) to visualize the relative importance of a given attribute in that class by dividing the difference in utility between highest and lowest level for a single attribute by the sum of the differences of all attributes for that class, taking interaction effects into account [36]. An attribute with an IS of 1 represents the most important attribute, while an attribute with an IS of 5 represents the least important

attribute. Furthermore, we also calculated overall importance scores, by taking class probability into account.

Expected uptake of the vaccine

Choice probabilities (mean uptakes) were calculated to provide a way to convey DCE results to policy makers that are more easily understandable. We calculated the choice probability (i.e. the mean uptake) for a base case vaccination for three given outbreaks by taking the exponent of the total utility for vaccination divided by the exponent of utility of both vaccination and no vaccination taking the class probabilities into account. The base-case vaccination program was chosen to resemble real life situations, and included the following attribute levels: vaccine effectiveness 70%, supposed to be a safe vaccine, advised by friends and positive traditional media attention, and no out-of-pocket costs. Outbreaks were defined as mild, moderate and severe (respectively a susceptibility and severity of 5% and 5%; 10% and 25%; and 20% 75%).

Trade-offs

We calculated willingness-to-pay (WTP) values for the effectiveness of the vaccine attribute for mild, moderate and severe outbreaks (respectively a susceptibility and severity of 5% and 5%; 10% and 25%; and 20% and 75%). A WTP value represents how much one is willing to pay for a one unit change in the attribute of interest, and is calculated by taking the ratio of the derivative of the effectiveness attribute and the derivative of out-of-pocket costs. Since effectiveness was included as both a main effect and as part of an interaction effect with susceptibility to the disease and severity of the disease, it is necessary to calculate the derivatives with respect to all parts of the utility function where the attribute appears [37]. Because a latent class model was used, overall WTP measures can be calculated by weighing the conditional WTP values by the probability that respondents belong to a given class. We computed the confidence intervals using the Krinsky and Robb procedure [38] (additional file 3).

We used NLogit 4.0 software (www.limdep.com) to estimate the latent class models and SPSS 21.0 software (http://www-01.ibm.com/software/analytics/spss/) for all other analysis.

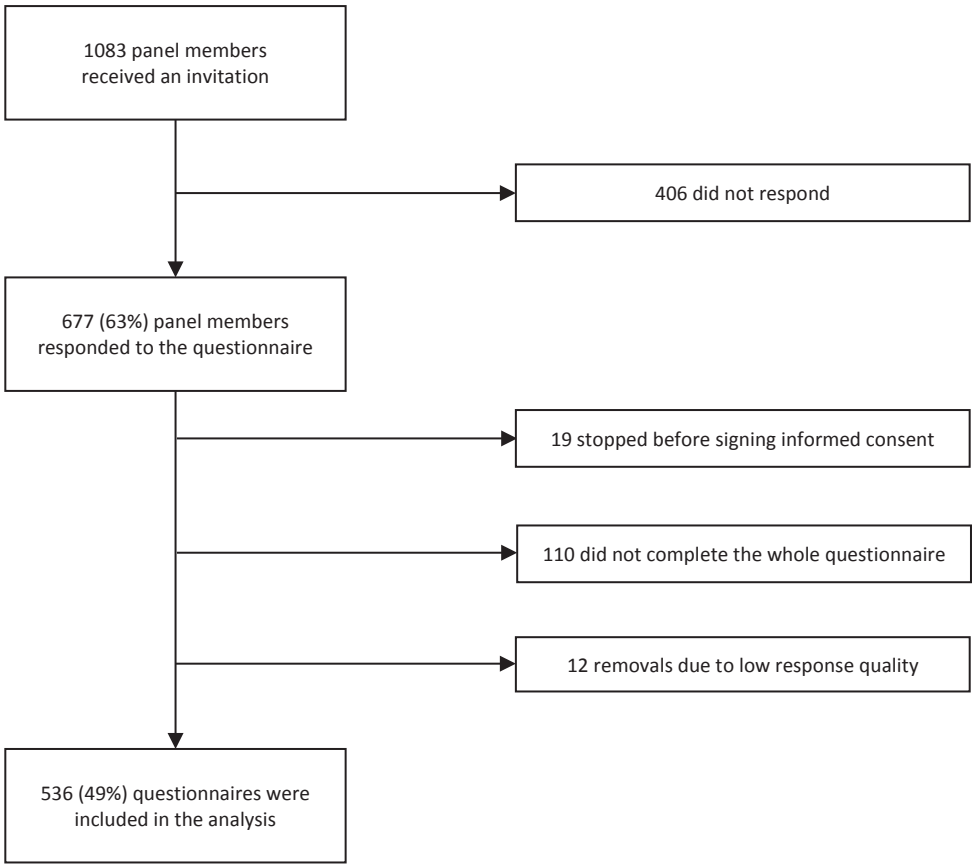
RESULTS

Respondents

The participation rate was 677/1083 (63%, Figure 1), which reflected the expected response rate for this online panel. Of the 677 respondents, 548 completed the questionnaire. Respondents who completed the questionnaire did not differ regarding sex ($p = 0.11$) or educational level ($p = 0.11$) compared to respondents who did not complete the questionnaire.

However, respondents who completed the questionnaire were younger (median age 50 vs. 53, p,0.01). Twelve respondents were excluded from the analysis, because they completed the questionnaire too quick; they completed the whole questionnaire in less than five minutes. Data of 536 (49%) respondents were included in the analysis.

Figure 1 | Overview of respondents accessing the study.



Respondents had a median age of 50 years (interquartile range (IQR): 35–64), with a minimum of 18 and a maximum age of 89 years old (Table 2). 30% had a high educational level and 22% of the respondents indicated that they had a positive attitude regarding vaccination, i.e. that they would always get vaccinated. The sample was representative for the Dutch population regarding age, gender, educational level and region.

Table 2 | Characteristics of respondents who completed the DCE survey (N=536).

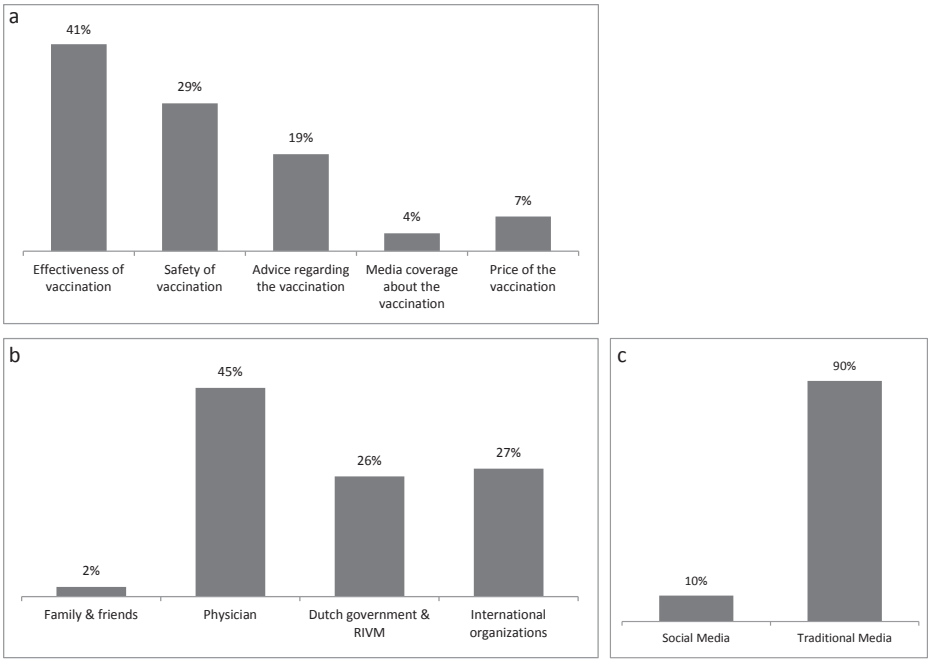
Characteristics	Subcategory	Sample statistics		CBS statistics
		Median	*IQR*	
Age in years		50	35-64	
		n	*%*	*%*
Age groups	18-24 years	49	9.2	11
	25-34 years	78	15	16
	35-44 years	84	16	19
	45-54 years	107	20	19
	55-64 years	92	17	16
	> 65 years	126	24	19
Gender	Male	289	54	49
Country of birth	Netherlands	517	96	-
Educational level	Low	184	34	34
	Average	192	36	40
	High	160	30	26
Civil status	Married	296	55	-
	Registered partnership	48	9.0	-
	Unmarried	133	25	-
	Divorced	38	7.1	-
	Widow / widower	21	3.9	-
Children	Yes	345	64	-
Income in euros per year	Minimal (< 11.000)	37	6.9	-
	Less than modal (11.000-23.000)	69	13	-
	Modal (23.000-34.000)	127	24	-
	1-2 times modal (34.000-56.000)	103	19	-
	2 times modal or more (>56.000)	78	15	-
	Do not know / do not want to say	122	23	-
Religion	Yes	244	46	-
Perception of health	Lower health than average	41	7.6	-
	Medium health	195	36	-
	Better health than average	300	56	-
Attitude regarding vaccination	Always get vaccinated	120	22	-
	Only if benefits > harms	259	48	-
	Only if benefits > harms, but I do not think this is the case in the real world	116	22	-
	Never get vaccinated, even if benefits > harms	41	7.6	-
Belongs to target group for seasonal flu vaccine	Yes	239	45	-
	No	270	45	-
	No, but receives flu vaccination via work	27	5.0	-
Belongs to the target group and received seasonal flu vaccination last year	Yes	160	60	-

Abbreviations: *CBS* Statistics Netherlands; *IQR* Interquartile range.

The median completion time of the whole questionnaire was 13 minutes (median, IQR: 9.6–19). It took respondents a median of 6.3 minutes (IQR: 4.2–9.6) to complete 16 choice tasks. The time respondents needed to fill in one choice task decreased from a median of 39 seconds (IQR: 20–62) for choice task 1 to 15 seconds (IQR: 10–24) for choice task 16. 67% of the respondents marked the number of choice tasks as 'exactly the good number' and 76% marked the questions as clear or very clear. A minority (13%) of the respondents found the questions hard or very hard to answer. Most of the respondents found the topic interesting or very interesting (87%). Responses to the free text question indicated that respondents felt that they were adequately informed to answer the questions in the questionnaire.

Direct ranking showed that respondents considered effectiveness the most important vaccine specific attribute, followed by safety of the vaccine and advice regarding the vaccine (Figure 2a–2c). Respondents marked their doctors' advice as most important, followed by the advice of international organizations and the advice of the Dutch government & National Institute for Public Health and the Environment (Dutch abbreviation: RIVM). Traditional media influenced the decision regarding vaccination more than social media.

Figure 2 | a. Direct ranking of attributes[1]; b. Direct ranking of attribute levels[2]; c. Direct ranking of attributes levels[3]



Notes: [1]The percentages represent the proportion of people that ranked that vaccination program attribute as most important when deciding on vaccination. [2]The percentages represent the proportion of people that ranked that vaccination program attribute level of advice regarding the vaccination as most important when deciding on vaccination. [3]The percentages represent the proportion of people that ranked that vaccination program attribute level of media coverage about the vaccination as most important when deciding on vaccination.

Discrete choice experiment results

The 'no vaccination' option was chosen in 37% of the choice sets. 61 respondents (11.0%) always chose the 'no vaccination' option. 113 respondents (21%) never chose the 'no vaccination' option.

Using a latent class model, two classes were identified (Table 3). The average class probabilities within the sample were 0.63 for class 1 and 0.37 for class 2. The probability to belong to a specific class was dependent on two socio-demographic variables: the sex of the respondent and the attitude of the respondent regarding vaccination. Males and individuals who stated that they (possibly) wanted to get vaccinated had the highest chance to belong to latent class 1, while females and individuals who stated that they would never get vaccinated had the highest chance to belong to latent class 2. Other socio-demographic variables were not significantly explaining class assignment probabilities.

The sign of the coefficient indicates whether the attribute had a positive or negative effect on utility (Table 3). For example, the positive sign for effectiveness and for side effects unknown, but expected to be safe indicated that an effective and safe vaccination was preferred over a vaccination which was less effective and with which there was no experience yet. The negative sign for out-of-pocket costs of vaccination indicated that respondents preferred vaccinations with lower out-of-pocket costs. The positive sign of the constant indicates that, everything else being equal, respondents preferred no vaccination over vaccination.

Nearly all of the vaccine specific characteristics were statistically significant (Table 3), proving to influence respondents' preference for vaccination. The interactions between the disease specific characteristics and effectiveness were significant and positive. This indicates that the preference for the level of effectiveness of a vaccination is dependent upon the levels of severity and susceptibility. If the susceptibility to or severity of a disease are higher, while the effectiveness of a vaccination is the same, preference for vaccination increases relative to no vaccination. In other words, if the susceptibility to the disease or the severity of a disease is higher, lower vaccination effectiveness will result in the same utility level. Note that the two disease characteristics cannot be included as a main effect but only as interaction effects, since they are scenario variables that are constant across all vaccine alternatives. All other 2-way interactions were not statistically significant.

When comparing the overall importance scores (Table 3) with the direct ranking question (Figure 2a), effectiveness of the vaccine was considered the most important attribute in both preference elicitation methods, especially when an outbreak was more serious, and media coverage of the vaccine as the least important attribute. These results support the convergent validity of the results.

Table 3 | Preferences of respondents for vaccinations in pandemic situations based on a panel latent class logit model (N=536).

Attributes		Latent class 1			Latent class 2			Overall
		Value[2]	*S.E.*	*IS*[3]	*Value*[2]	*S.E.*	*IS*[3]	*IS*[3]
Constant (no vaccination)		0.22**	0.05		2.46**	0.15		
Effectiveness of vaccine (per 10%)		0.07**	0.01	1	0.01	0.03	1	1
Side effects unknown, but expected to be safe (ref)[1]		0.21	-	4	0.27	-	5	5
Side effects unknown, no experience yet		-0.21**	0.01		-0.27**	0.05		
Family and/or friends recommend (ref)[1]		-0.33	-	2	-0.04	-	3	2
Family and/or friends discourage		-0.50**	0.03		-0.39**	0.11		
Your doctor recommends		0.42**	0.03		0.51**	0.10		
Your doctor discourages		-0.67**	0.03		-1.02**	0.15		
Dutch government & RIVM recommend		0.58**	0.03		0.48**	0.10		
International organizations recommend		0.49**	0.03		0.45**	0.09		
Traditional media is positive (ref)[1]		0.15	-	5	0.39	-	4	4
Traditional media is negative		-0.20**	0.03		-0.28**	0.09		
Social / interactive media is positive		0.16**	0.02		0.18*	0.08		
Social / interactive media is negative		-0.11**	0.03		-0.29**	0.09		
Out-of-pocket costs of the vaccination (per 10 euro)		-0.06**	0.00	3	-0.17**	0.01	2	3
Interactions between attribute levels								
Interaction: effectiveness of vaccine (per 10%) x susceptibility to the disease (per 10%)		0.12**	0.00	-	0.11**	0.00	-	-
Interaction: effectiveness of vaccine (per 10%) x severity of the disease (per 10%)		0.02**	0.00	-	0.01**	0.00	-	-
Class probability model	**Subcategory**							
Constant		-	-	-	1.63**	0.47	-	-
Sex	Male	-	-	-	- 0.80**	0.22	-	-
Attitude regarding vaccination	I will never get vaccinated (ref)	-	-	-	<0.01	<0.01	-	-
	I will always get vaccinated	-	-	-	-4.09**	0.73	-	-
	I will only get vaccinated if advantages > disadvantages	-	-	-	-1.88**	0.50	-	-
	I will only get vaccinated if advantages > disadvantages, however I do not think that is the case in the real world	-	-	-	-0.78	0.48	-	-
Class probability								
Average		0.63	-	-	0.37	-	-	-
Model fit								
AIC		1.64	-	-	-	-	-	-
Log likelihood		-6989	-	-	-	-	-	-
R^2		0.26	-	-	-	-	-	-

Abbreviations: S.E. = standard error; IS = importance score; REF = reference level; RIVM = Dutch abbreviation for National Institute for Public Health and the Environment.

Notes: Effects coded variables used for safety of the vaccine, advice about the vaccine, media coverage about the vaccine. Number of observations: 25728 (16*3*536). [1]The values of the vaccination program attributes reference levels equals the negative sum of the coefficients of the included attribute. [2]**, * denotes significance at the 1% and 5% respectively. [3]The IS were calculated for a severe outbreak with a susceptibility to the disease of 20% and a severity of the disease of 75%.

Preference heterogeneity was substantial; respondents belonging to latent class 1 seemed to place more weight on the effectiveness of the vaccine than respondents of latent class 2 (IS of 2 for class 1 compared to an IS of 5 for class 2, in case of a mild outbreak). However, in case of a severe outbreak, effectiveness was the most important attribute for both latent classes. Respondents belonging to class 2 were more influenced by the media and more sensitive to costs than respondents belonging to latent class 1 (respectively an IS of 3 and 1 for class 2, and an IS of 5 and 3 for class 1, in case of a mild outbreak of the disease). For respondents of both classes, the advice regarding vaccination of others was important. Respondents in class 1 were most influenced by the advice of the government & RIVM and international organizations, while respondents in latent class 2 were most influenced by the recommendation or discouraging of their physician.

Trade-offs

Based on the expressed preferences, respondents were willing to pay €6.0 (95% Confidence Interval: €3.7-€8.3) to receive a 10% more effective vaccine in case of a mild pandemic outbreak (Table 4). If a pandemic outbreak was more severe the willingness to pay for a vaccine which was 10% more effective increased up to €20 (€18-€22) in case of a moderate outbreak and €39 (€36-€44) in case of a severe outbreak.

Table 4 | Willingness to pay.

Attribute	To receive a vaccination	WTP (€, CI)		
		Mild pandemic[1]	Moderate pandemic[2]	Severe pandemic[3]
Effectiveness of vaccine	With 10% more effectiveness	6.0 (3.7-8.3)	20 (18-22)	39 (36-44)

Abbreviations: WTP = willingness to pay; € = euro; CI = 95% confidence interval based on the Krinsky and Robb method adjusted for class probabilities and taking into account interaction effects (see Figure S3 for more information).
Notes: [1]Mild pandemic is defined as a disease with a susceptibility of 5% and a severity of 5%. [2]Moderate pandemic is defined as a disease with a susceptibility of 10% and a severity of 25%. [3]Severe pandemic is defined as a disease with a susceptibility of 20% and a severity of 75%.

Expected uptake of the vaccine

The mean predicted uptake of the base-case vaccination program increased from 50% in a mild pandemic up to 88% for a severe pandemic (Figure 3a-3c). The more serious an outbreak was, the more the predicted uptake depended on effectiveness of a vaccine, e.g. a vaccine that was 40% less effective compared to the base case vaccination decreased the vaccination uptake 11, 20 and 28 percent points, for a mild, moderate or severe outbreak respectively. Irrespective of the disease scenario; higher out-of- pocket costs had a relatively large impact on the vaccination uptake, compared to the base case vaccination which was free.

Figure 3 | a. Estimates for predicted probability of participation; values for a mild outbreak[1,2,3,4]; b. Estimates for predicted probability of participation; values for a moderate outbreak[1,2,5,6]; c. Estimates for predicted probability of participation; values for a severe outbreak[1,2,7,8].

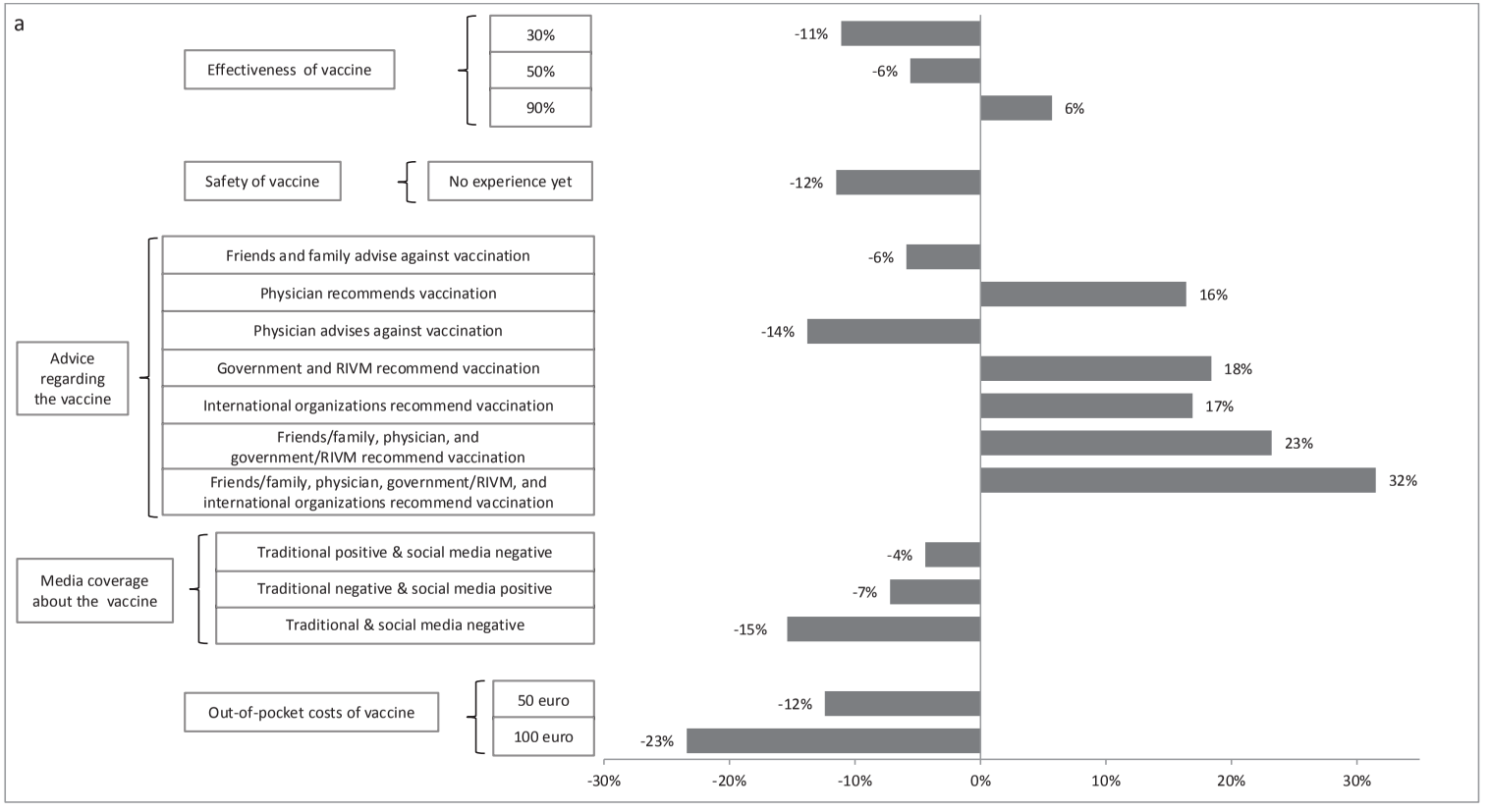


Figure 3 | Continued

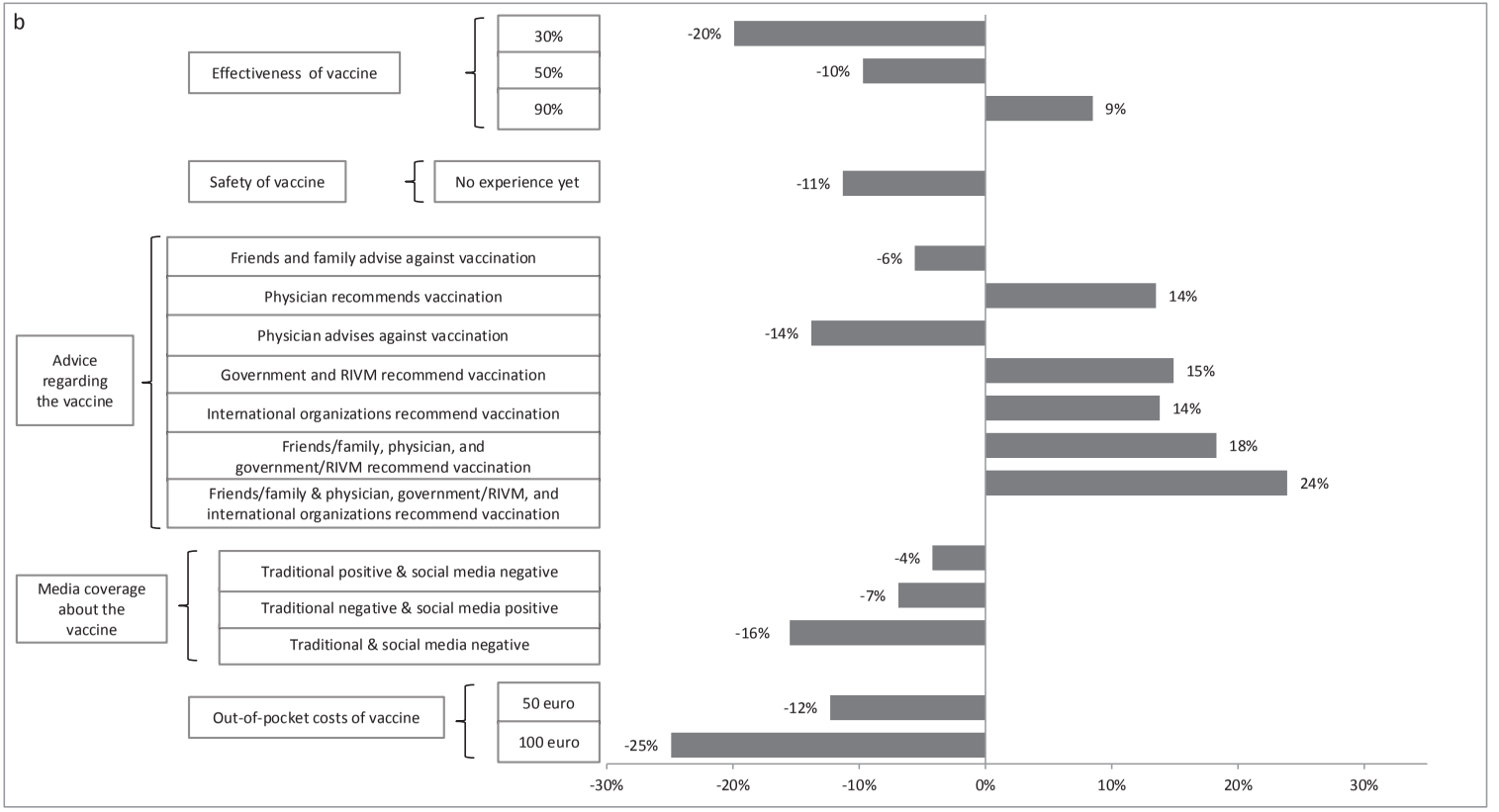
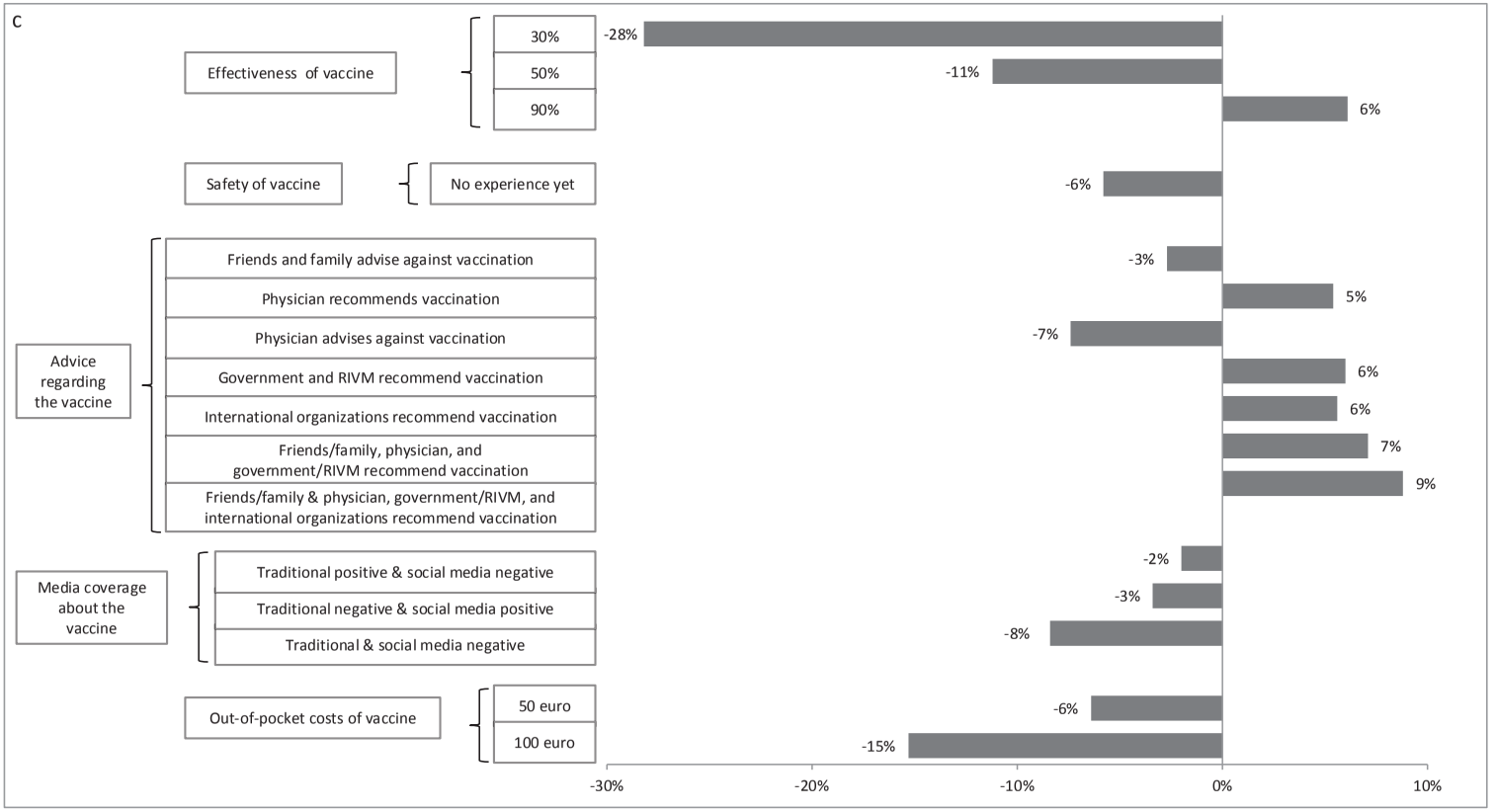


Figure 3 | Continued



Notes: [1]The percentages represent the change in probability compared to base case vaccination. [2]The base case vaccination is 70% effective, supposed to be safe, recommended by friends/family, the traditional media is positive and there are no out-of-pocket costs. This base case is indicated as zero change in the probability of the x-axis. [3]A mild outbreak is defined as 5% of the population getting sick and 5% of the population getting severe symptoms. [4]Probability of base case vaccination in this scenario = 50%. [5]A moderate outbreak is defined as 10% of the population getting sick and 25% of the population getting severe symptoms. [6]Probability of base case vaccination in this scenario = 65%. [7]A severe outbreak is defined as 20% of the population getting sick and 75% of the population getting severe symptoms. [8]Probability of base case vaccination in this scenario = 88%.

Furthermore, recommendation of the vaccine by physicians, the government & RIVM or international organizations resulted in a substantial increase of the predicted uptake of the base case program (e.g. an increase of 16, 18 and 17 percent points respectively in case of a mild outbreak). Assuming that all bodies advised positively regarding the vaccine (including friends and family) the predicted uptake increased with 32 percent points in case of a mild outbreak.

DISCUSSION

This DCE showed that effectiveness, safety and out-of-pocket costs of the vaccine, as well as advice regarding and media coverage about the vaccine all influenced the general populations' preference for pandemic vaccinations. Preference heterogeneity was substantial; two latent classes with different preferences were identified by a latent class model. Female respondents and individuals who stated that they would never get vaccinated were more influenced by the media and more sensitive to costs than male respondents and individuals who stated that they were (possibly) willing to get vaccinated. As expected, respondents preferred and were willing to pay more for more effective vaccines, especially if the outbreak was more serious. Changes in effectiveness, out-of-pocket costs of the vaccine and in the body that advises the vaccine substantially influenced the predicted uptake.

This is the first DCE investigating how characteristics of pandemic vaccinations influence preferences for vaccination programs in different pandemic outbreaks. Two systematic reviews assessed which factors were associated with uptake of the Influenza A (H1N1) pandemic vaccine. These also showed the need for targeted messaging to reach vaccination goals [7,8]. Especially the conclusion of one of these reviews [7] that social pressure and confidence in sources of information had an effect on the intention to vaccinate, is in line with our results. To gain insight in factors explaining willingness to vaccinate against Influenza A(H1N1) in The Netherlands, a questionnaire study was conducted among the general Dutch population during the 2009–2010 pandemic [39]. Similar results as we found were reported: people who were afraid of the disease, who perceived it as a severe disease, who believed in the efficacy of the vaccine and who trusted the information the government provided had higher odds for vaccination. Furthermore, the majority of respondents trusted the information provided by their general practitioner and more than half of the respondents trusted the information provided by the Dutch government and RIVM. Another questionnaire study regarding the Influenza A(H1N1) pandemic in the Netherlands showed results that are in line with our results as well; most respondents wanted to receive information about infection prevention from municipal health services, health care providers, and the media. Higher levels of intention to receive vaccination were associated with increased government trust, fear or worry about the disease, and perceived vulnerability to the disease [40].

Several DCEs on non-pandemic vaccines [12,13,41] showed the influence of similar characteristics on vaccination preferences as we found in our study. In a DCE on preferences for HPV vaccination [12], it was found that the degree of protection positively influenced the preference of girls for vaccination, while the risk of side effects had a negative effect. A DCE among parents preferences for influenza vaccination for their children [13], showed that the efficacy of a vaccination and the recommendation of physicians positively influenced parents' preferences, while the risk of temporary side effects had a negative effect. A DCE on marginal WTP for HIV vaccines [41] found that biomedical characteristics of a hypothetical HIV vaccine, such as efficacy, vaccine induced seropositivity and side effects, were the most important attributes for vaccination programs.

Our results suggest that side effects of the vaccine are less important than the other included attributes when deciding on vaccinations, while in other studies (including DCEs) safety of vaccinations was dominant [7,12,13,41,42]. This difference can probably be assigned to the choice of attribute levels since respondents in the current DCE were informed that the chance of side effects was expected to be low and either comparable to vaccines that are already on the market or expected to be low, but with no experience with a similar vaccination yet, i.e. a totally new vaccination. Our study showed that preference heterogeneity was substantial. Findings on heterogeneity are supported by a focus group study on acceptance of hypothetical pandemic vaccinations in Canada, where parents with non-mainstream beliefs showed different concerns regarding vaccinations [43].

This study had several limitations. First, we measured preferences for hypothetical vaccines in hypothetical pandemic outbreaks. Although we were not able to measure the external validity of our results, the results may be very helpful in helping to prepare for pandemic outbreaks. Additionally, the signs of the coefficients were generally consistent with our a priori hypothesis (a higher susceptibility to the disease, a higher severity of the disease and a higher effectiveness would have a positive effect on vaccination) and therefore, theoretically valid. Second, the participation rate of 49% was not optimal and selection bias cannot be excluded. However, this participation rate is equal or even higher than most other DCEs in health care. Furthermore, the participation rate was comparable to the average rate of the internet panel we used. We expect our results to be generalizable since age, gender, level of education and region of our sample are comparable to that of the general population of the Netherlands. Third, due to both the number and the type of attributes and levels that respondents needed to take into account when completing the choice tasks, it can be expected that respondents might have experienced difficulties, which might have influenced the results. However, piloting and think-a-loud interviews in the preparation phase, as well as questions that assessed the experienced difficulty of the questionnaire showed that the majority of respondents had no problems with completing the tasks. Fourth, we included safety of the vaccine as a categorical attribute, instead of a numerical

attribute, which would have helped respondents to compare risks of vaccinations with risks of the disease. However, when designing the DCE, expert interviews showed that safety of the vaccine is more or less a fixed attribute (either being sure that the vaccine will be safe, or that there is no experience with the vaccine yet and there is thus a chance of long-term side effects). Therefore, we included the safety of the vaccine as an attribute with categorical levels.

Insights in the factors influencing the intention to accept or decline a pandemic vaccine may have implications for both national and international policy and for further research. When communicating public health messages regarding vaccination, one should be aware of preference heterogeneity and therefore use different sources and channels to distribute the messages [44]. The current study provides guidance on how to target public health messages, by the identification of two classes with different preferences for pandemic vaccinations. To immediately reduce the number of susceptible people, a possible strategy could be to target the message for the first phase of a vaccination program to the more vaccination minded persons, here latent class 1. This can be done by using the government and RIVM as bodies to advice the vaccine to males and focus more on the expected effectiveness of the vaccine. Next, physicians can advise females to take the vaccine. Additionally, out-of-pocket costs need to be as low as possible, as our study showed the negative relation between out-of- pocket costs and vaccination decisions. For public health messages during vaccination programs, it is also important to monitor side effects. Updates of the side effects of the vaccine need to be given on a regularly basis to make sure that an informed choice can be made and to reduce fear of the side effects of the vaccine. Furthermore, policy makers can use the expected uptake probability of hypothetical vaccinations when predicting the number of vaccinations that is needed. Although these numbers are rough estimates and it is not known if they are externally valid, the expected uptake can still be useful when other information is lacking. Additionally, these numbers can guide communication on the expected vaccination uptake. Since this is the first quantitative study in motivations for pandemic vaccinations, we do not know to what extent differences exist between countries regarding preferences for vaccinations. There is some evidence, including a questionnaire study in four countries investigating reasons why high risk people reject influenza vaccination in four countries of Europe suggests differences between respondents of the different countries [45]. Therefore, further international research is recommended.

We conclude that various disease and vaccination program attributes influence respondents' preferences for pandemic vaccination programs. Agencies responsible for preventive measures during pandemics can use the findings of this study that out-of-pocket costs and the way advice is given affect vaccination uptake to change the way vaccination is marketed during future pandemic outbreaks. The preference heterogeneity shows that information regarding vaccination needs to be targeted differently depending on gender and willingness to get vaccinated.

REFERENCES

1. Viboud C, Miller M, Olson D, et al. Preliminary Estimates of Mortality and Years of Life Lost Associated with the 2009 A/H1N1 Pandemic in the US and Comparison with Past Influenza Seasons. PLoS Curr. 2010;2:RRN1153.
2. Nguyen-Van-Tam JS, Hampson AW. The epidemiology and clinical impact of pandemic influenza. Vaccine. 2003;21:1762–8.
3. Murray CJ, Lopez AD, Chin B, et al. Estimation of potential global pandemic influenza mortality on the basis of vital registry data from the 1918–20 pandemic: a quantitative analysis. Lancet. 2006;368:2211–8.
4. Ferguson NM, Cummings DA, Cauchemez S, et al. Strategies for containing an emerging influenza pandemic in Southeast Asia. Nature. 2005;437:209–14.
5. Longini IM Jr, Nizam A, Xu S, et al. Containing pandemic influenza at the source. Science. 2005;309:1083–7.
6. Stern EK, Young S, Amlôt R, et al. Assessment Report on EU-wide Pandemic Vaccine Strategies. European Commission; 2010. Available from: http://ec.europa.eu/health/communicable_diseases/docs/assessment_vaccine_en.pdf.
7. Bish A, Yardley L, Nicoll A, et al. Factors associated with uptake of vaccination against pandemic influenza: a systematic review. Vaccine. 2011;29:6472–84.
8. Nguyen T, Henningsen KH, Brehaut JC, et al. Acceptance of a pandemic influenza vaccine: a systematic review of surveys of the general public. Infect Drug Resist. 2011;4:197–207.
9. Mereckiene J, Cotter S, Weber JT, et al. Influenza A(H1N1)pdm09 vaccination policies and coverage in Europe. Euro Surveill. 2012;17:pii=20064.
10. De Bekker-Grob EW, Ryan M, Gerard K. Discrete choice experiments in health economics: a review of the literature. Health Econ. 2012;21:145–72.
11. Ryan M, Gerard K. Using discrete choice experiments to value health care programmes: current practice and future research reflections. Appl Health Econ Health Policy. 2003;2:55–64.
12. De Bekker-Grob EW, Hofman R, Donkers B, et al. Girls' preferences for HPV vaccination: a discrete choice experiment. Vaccine. 2010;28:6692–7.
13. Flood EM, Ryan KJ, Rousculp MD, et al. Parent preferences for pediatric influenza vaccine attributes. Clin Pediatr. 2011;50:338–47.
14. Ryan M. Discrete choice experiments in health care. BMJ. 2004;328:360–1.
15. Hensher DA, Rose JM, Greene WH. Applied choice analysis: a primer. New York: Cambridge University Press; 2005.
16. Lancaster KJ. A new approach to consumer theory. J Polit Econ. 1966;74:132–57.
17. McFadden D. Conditional logit analysis of qualitative choice behavior. In: Zarembka P, editor. Frontiers of Economics. New York: Academic Press; 1974. p. 105-42.
18. Bridges JF, Hauber AB, Marshall D, et al. Conjoint analysis applications in health–a checklist: a report of the ISPOR Good Research Practices for Conjoint Analysis Task Force. Value Health. 2011;14:403–13.
19. Lancsar E, Louviere J. Conducting discrete choice experiments to inform healthcare decision making: a user's guide. Pharmacoeconomics. 2008;26:661–77.
20. Coast J, Al-Janabi H, Sutton EJ, et al. Using qualitative methods for attribute development for discrete choice experiments: issues and recommendations. Health Econ. 2012;21:730–41.
21. Janz NK, Becker MH. The Health Belief Model: a decade later. Health Educ Q. 194;11:1–47.
22. Rogers RW. A protection motivation theory of fear appeals and attitude change. JRL. 1975;91:93–114.
23. Coe AB, Gatewood SB, Moczygemba LR, et al. The use of the health belief model to assess predictors of intent to receive the novel (2009) H1N1 influenza vaccine. Inov Pharm. 2012;3:1–11.
24. Teitler-Regev S, Shahrabani S, Benzion U. Factors Affecting Intention among Students to Be Vaccinated against A/H1N1 Influenza: A Health Belief Model Approach. Adv Prev Med. 2011;2011:353207.

25. Reed Johnson F, Lancsar E, Marshall D, et al. Constructing experimental designs for discrete-choice experiments: report of the ISPOR Conjoint Analysis Experimental Design Good Research Practices Task Force. Value Health. 2012;16:3–13.
26. Hall J, Fiebig DG, King MT, et al. What influences participation in genetic carrier testing? Results from a discrete choice experiment. J Health Econ. 2006;25:520–37.
27. Peters E, Vastfjall D, Slovic P, et al. Numeracy and decision making. Psychol Sci. 2006;17:407–13.
28. Howard K, Salkeld G. Does attribute framing in discrete choice experiments influence willingness to pay? Results from a discrete choice experiment in screening for colorectal cancer. Value Health. 2009;12:354–63.
29. Tversky A, Kahneman D. The framing of decisions and the psychology of choice. Science. 1981;211:453–8.
30. Kjaer T, Bech M, Gyrd-Hansen D, et al. Ordering effect and price sensitivity in discrete choice experiments: need we worry? Health Econ. 2006;15:1217–28.
31. Hak T, van der Veer K, Jansen H. The Three-Step Test-Interview (TSTI): An observation-based method for pretesting self-completion questionnaires. Survey Research Methods. 2008;2:143–50.
32. Marshall D, Bridges JF, Hauber AB, et al. Conjoint Analysis Applications in Health - How are Studies being Designed and Reported?: An Update on Current Practice in the Published Literature between 2005 and 2008. Patient. 2010;3:249–56.
33. Orme BK. Getting started with Conjoint Analysis: Strategies for Product Design and Pricing Research. Madison, WI: Research Publishers LLC; 2006.
34. De Bekker-Grob EW, Rose JM, Donkers B, et al. Men's preferences for prostate cancer screening: a discrete choice experiment. Br J Cancer. 2013;108:533–41.
35. Swait J. A structural equation model of latent segmentation and product choice for cross-sectional revealed preference choice data. J Retailing Consum Serv. 1994;1:7–89.
36. Louviere JH, Hensher D, Swait JD. Stated choice models: analysis and application. Cambridge: Cambridge University Press; 2000.
37. De Bekker-Grob EW, Rose JM, Bliemer MC. A Closer Look at Decision and Analyst Error by Including Nonlinearities in Discrete Choice Models: Implications on Willingness-to-Pay Estimates Derived from Discrete Choice Data in Healthcare. Pharmacoeconomics. 2013;31:1169–83.
38. Krinsky I, Robb AL. On Approximating the Statistical Properties of Elasticities. Rev Econ Stat 1986;68:715–9.
39. Bults M, Beaujean DJ, de Zwart O, et al. [Mexican flu: risk perception in the general public, precautionary measures and trust in information provided by the government]. Mexicaanse griep: risicoperceptie bij de bevolking, eigen maatregelen en vertrouwen in overheidsinformatie. Ned Tijdschr Geneeskd. 2010;154:A1686. *In Dutch.*
40. Van der Weerd W, Timmermans DR, Beaujean DJ,et al. Monitoring the level of government trust, risk perception and intention of the general public to adopt protective measures during the influenza A (H1N1) pandemic in The Netherlands. BMC Public Health. 2011;11:575.
41. Cameron MP, Newman PA, Roungprakhon S, et al. The marginal willingness-to-pay for attributes of a hypothetical HIV vaccine. Vaccine. 2013;31:3712–7.
42. Sadique MZ, Devlin N, Edmunds WJ, et al. The effect of perceived risks on the demand for vaccination: results from a discrete choice experiment. PLoS ONE. 2013;8:e54149.
43. Henrich N, Holmes B. The public's acceptance of novel vaccines during a pandemic: a focus group study and its application to influenza H1N1. J Emerg Health Threats. 2009;2:e8.
44. Freimuth V, Linnan HW, Potter P. Communicating the threat of emerging infections to the public. Emerg Infect Dis. 2000;6:337–47.
45. Kroneman M, van Essen GA, Paget WJ. Influenza vaccination coverage and reasons to refrain among high-risk persons in four European countries. Vaccine. 2006;24:622–8.

ADDITIONAL FILE 1: HYPOTHETICAL SCENARIO

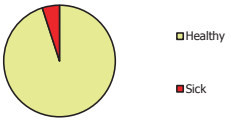
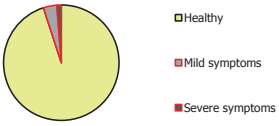
This scenario was presented to respondents in Dutch.

"Imagine, a new disease emerged abroad. This disease is highly contagious, because the disease spreads by droplets. All over the world, people become infected, also in the Netherlands. You can do some things yourself to prevent you of getting the disease (such as washing hands, etc.). However, the most effective preventive measure is a vaccination. Vaccinations are available for everyone in the Netherlands."

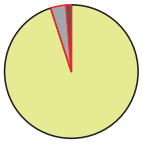
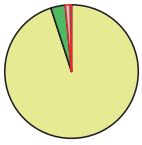
ADDITIONAL FILE 2: EXAMPLE OF A CHOICE SET

Choice sets were presented to respondents in Dutch.

Additional information regarding the new disease (the situation without vaccination):

Susceptibility to the disease	**50 out of 1.000** people will get sick (develop symptoms) 
Severity of the disease	**25% of all sick people** will get severe symptoms 

You can choose the following 3 options, what do you choose?

	Option 1: **No vaccination**	**Option 2:** **Vaccination A**	**Option 3:** **Vaccination B**
Effectiveness of the vaccine	**n.a.** This means: **50 out of 1.000** people will get sick  **25% of these sick people** will get severe symptoms	Effectiveness: **30%** This means: **15 out of 1.000** people will not get sick, due to vaccination **35 out of 1.000** people will get sick, despite the vaccination  **25% of these sick people** will get severe symptoms	Effectiveness: **70%** This means: **35 out of 1.000** people will not get sick, due to vaccination **15 out of 1.000** people will get sick, despite the vaccination  **25% of these sick people** will get severe symptoms
Safety of the vaccine **(long term severe side effects)**	No side effects	Unknown, but expected to be safe	Unknown, no experience yet
Advice about the vaccine	n.a.	Recommended by your doctor	Recommended by family and friends
Media coverage about the vaccine	n.a.	Radio, newspapers and television positive	Radio, newspapers and television negative
Out-of-pocket costs	0 euro	50 euros	50 euros

What do you choose for yourself? (please tick one box only)	☐	☐	☐

ADDITIONAL FILE 3: UTILITY FUNCTIONS

The utility function for no vaccination in the model with the best fit was:

$$V(\text{no vaccination})_{nsj|c} = \beta_{0|c} \quad \text{(Eq.1)}$$

The utility function for vaccination in the model with the best fit was:

$$V(\text{vaccination})_{nsj|c} = \beta_{1|c}\,\text{effectiveness}_{nsj|c} + \beta_{2|c}\,\text{side effects (unknown)}_{nsj|c} + \beta_{3|c}\,\text{advice (friends advise against)}_{nsj|c} + \beta_{4|c}\,\text{advice (doctor recommends)}_{nsj|c} + \beta_{5|c}\,\text{advice (doctor advises against)}_{nsj|c} + \beta_{6|c}\,\text{advice (government \& RIVM recommend)}_{nsj|c} + \beta_{7|c}\,\text{advice (international organizations recommend)}_{nsj|c} + \beta_{81|c}\,\text{media coverage (traditional media negative)}_{nsj|c} + \beta_{9|c}\,\text{media coverage (social media positive)}_{nsj|c} + \beta_{10|c}\,\text{media coverage (social media negative)}_{nsj|c} + \beta_{11|c}\,\text{costs}_{nsj|c} + \beta_{12|c}\,\text{effectiveness}_{nsj|c} \times \text{susceptibility}_{nsj|c} + \beta_{13|c}\,\text{effectiveness}_{nsj|c} \times \text{severity}_{nsj|c} \quad \text{(Eq.2)}$$

where

- $V_{nsj|c}$ represents the observable utility that respondent *n* belonging to class segment *c* has for alternative *j* in choice set s for vaccination;
- $\beta_{0|c}$ is an alternative-specific constant reflecting respondents' preferences for no vaccination compared to receive a vaccination for a certain class;
- $\beta_{1\text{-}11|c}$ are class-specific main-effects regression coefficients of the attributes, indicating the relative weight individuals place on certain attribute levels.
- $\beta_{12\text{-}13|c}$ are class-specific two-way interaction effects (i.e. an effect where the influence of one attribute depends on the level of another attribute).

In addition to the utility function for vaccination (Eq. 1 and Eq. 2), the final model allowed for several covariates to enter into the class assignment model. The class assignment utility function for the final model was:

$$V_{nc} = \beta_{0|c} + \beta_{1|c}\,\text{gender}_{n} + \beta_{2|c}\,\text{attitude to vaccination (always positive)}_{n} + \beta_{3|c}\,\text{attitude to vaccination (in favor when advantages > disadvantages)}_{n} + \beta_{4|c}\,\text{attitude to vaccination (in favor when advantages > disadvantages, but don't thinks that's the case in the real world)}_{n} \quad \text{(Eq. 3)}$$

The WTP was calculated by taking the ratio of the parameter for the effectiveness attribute to the parameter related to out-of-pocket costs. Since effectiveness was included as both a main and an interaction effect, it was necessary to take also values of the attributes susceptibility to the disease and severity of the disease into account when calculating the WTP [1]. Because a latent class model was used, overall WTP measures could be calculated by weighing the conditional WTP values by the probability that respondents belong to a given class (given by the class assignment probability in Equation (1) and (2)). We calculated the WTP for three different pandemic outbreaks, using the values 5%, 10% and 20% for susceptibility and the values 5%, 25% and 75% for severity.

REFERENCE

1. De Bekker-Grob EW, Rose JM, Bliemer MC. A Closer Look at Decision and Analyst Error by Including Nonlinearities in Discrete Choice Models: Implications on Willingness-to-Pay Estimates Derived from Discrete Choice Data in Healthcare. Pharmacoeconomics. 2013;31:1169-83.

6

What health plans do people prefer?

The trade-off between premium and provider choice

Domino Determann
Mattijs S. Lambooij
Esther W. de Bekker-Grob
Arthur P. Hayen
Marco Varkevisser
Frederik T. Schut
G. Ardine de Wit

Accepted for publication in Social Science & Medicine. July 2016

ABSTRACT

Within a healthcare system with managed competition, health insurers are expected to act as prudent buyers of care on behalf of their customers. To fulfil this role adequately, understanding consumer preferences for health plan characteristics is of vital importance. Little is known, however, about these preferences and how they vary across consumers. Using a discrete choice experiment (DCE) we quantified trade-offs between basic health plan characteristics and analysed whether there are differences in preferences according to age, health status and income. We selected four health plan characteristics to be included in the DCE: (i) the level of provider choice and associated level of reimbursement, (ii) the primary focus of provider contracting (price, quality, social responsibility), (iii) the level of service benefits, and (iv) the monthly premium. This selection was based on a literature study, expert interviews and focus group discussions. The DCE consisted of 17 choice sets, each comprising two hypothetical health plan alternatives. A representative sample (n=533) of the Dutch adult population, based on age, gender and educational level, completed the online questionnaire during the annual open enrolment period for 2015. The final model with four latent classes showed that being able to choose a care provider freely was by far the most decisive characteristic for respondents aged over 45, those with chronic conditions, and those with a gross income over €3000/month. Monthly premium was the most important choice determinant for young, healthy, and lower income respondents. We conclude that it would be very unlikely for half of the sample to opt for health plans with restricted provider choice. However, a premium discount up to €15/month by restricted health plans might motivate especially younger, healthier, and less wealthy consumers to choose these plans.

INTRODUCTION

The health insurance system in the Netherlands was radically reformed by the introduction of a universal health insurance scheme with managed competition in 2006 [1, 2]. Similar reforms were introduced in other Western countries, e.g. Switzerland, Germany, Belgium, and the United States (Medicare Advantage, Medicare part D and ACA Marketplaces) [3, 4]. A key feature of the model of managed competition is that health insurers are provided with incentives to act as prudent buyers of care on behalf of their enrolees. To that end, insurers are expected to compete on the dimensions that consumers find important, for example price and quality of care, in order to maintain or increase their market share [5, 6]. Dutch consumers are expected to critically assess the available health plans annually, and to switch to another insurer's health plan if that better meets their preferences. Eventually, this competition in the health insurance market might result in better health care outcomes in terms of quality and costs.

The model of managed competition has, however, two related caveats. First, there is growing empirical evidence that many people are having trouble making optimal health plan choices because they face difficulties in understanding relevant health plan differences [7-10]. If consumers do not have the information and/or cognitive abilities to adequately choose an insurance plan, or when search costs and switching costs are substantial, competition among health insurers may not yield the desired outcomes. Such barriers to adequate decision making are likely to be a major cause of consumer inertia (or a preference for the status quo) typically observed in health insurance markets [11]. Second, given that consumers are having trouble making optimal health plan choices, market design and regulation may be needed to improve outcomes. As recently argued by Handel & Kolstad [10]: *"How consumers value different product attributes, what consumers know about those attributes, and how these preferences and information translate into choices is fundamental to market design and regulation, for health insurance and beyond. Without detailed knowledge of these micro-foundations it is difficult to precisely answer key policy questions such as which type of plans to allow insurers to offer and how those plans should be presented and priced."*

Actual health plan choices are widely studied. Research shows, for example, that younger consumers are more sensitive to premium compared to older consumers, who are more sensitive to quality [12-14]. A drawback of studying actual health plan choices is that only preferences that are observable from the available market data can be revealed. In addition, health insurance premiums are heavily regulated, which makes inferring willingness-to-pay from market observations difficult or even impossible. Most of these revealed preference studies do not provide insights in the relative importance of health plan characteristics. An exception is a recent study among Massachusetts state employees that examined to what extent enrolees of an insurance plan were willing to switch to a limited provider

network in return for a premium discount [15]. The findings suggest that switching to a limited provider network is very sensitive to financial incentives and that the healthiest individuals are somewhat more likely to switch, although the differences by health status are not substantial.

In this paper, we use the method of discrete choice experiments (DCEs) to elicit and quantify consumer preferences and trade-offs for health plan characteristics. A DCE is a stated-preference technique that has become increasingly popular in health economics over the past years [16]. This technique originates from mathematical psychology [17] and is grounded in random utility theory [18, 19]. In DCEs, respondents are presented with a series of hypothetical scenarios (called choice sets) in which they are asked to choose between two or more alternatives that are distinguished from each other by systematically varying characteristics (called attributes) [20, 21]. Several studies used DCEs, or other stated-preference techniques, to examine consumer preferences for health plans in various institutional contexts, such as community micro health insurance in Malawi [22] and health insurance settings in Colombia [23], the US [24, 25], and Switzerland [26]. In The Netherlands, several DCE-studies were performed just before [27, 28] and shortly after the Dutch health insurance reform [29, 30]. In addition, two other DCEs [31, 32] examined consumer sensitivity to various incentives to visit preferred providers. The most recent DCE study in this field stems from the year 2009 and was conducted among a small sample (n=97) of Dutch hospital patients only [33]. Recent studies among large representative samples of the national adult population are lacking.

The aim of this paper is twofold. First, we want to quantify consumers' trade-offs between basic health plan characteristics after almost ten years of experience with managed competition in the Dutch health care system. Since the reforms in 2006, additional changes have taken place, including more attention for quality of care, and an increase in selective contracting practices. In addition, people are more accustomed to the role they are expected to play in the system: being critical and choosing as consumers. They are faced with widespread annual media attention around the choice of health plan, and are now also used to experience the consequence of their choices made in the preceding year. In contrast to most previous DCE studies, our study is based on a large representative sample of the national adult population. Second, we want to analyse whether these consumer preferences differ according to age, income, and health status.

MATERIALS AND METHODS

Study context

Since the fundamental health system reform in 2006, Dutch citizens are obliged to buy individual basic health insurance with a legally standardized comprehensive basic benefit package at a community-rated premium from a private insurance company (for a more detailed description of the Dutch health care system, see [5, 34] and for comparisons with systems from other countries, see [35]). Insurers on the other hand are obliged to accept any applicant for basic health insurance, irrespective of the applicant's health. Additionally, a supplementary insurance that covers benefits not included in the basic benefit package (primarily dental care for adults and physiotherapy, for which no referral by a general practitioner is needed), can be taken out by citizens on a voluntarily basis in an unregulated market with risk-rated premiums. Dutch citizens have the possibility to switch health plans annually during a 2-month open enrolment period. In 2015, citizens could choose one of the 71 basic health plans offered by 25 insurers belonging to one of the in total 9 different health insurance companies [36]. Basic health plans may vary in terms of reimbursement – in-kind, in-cash or a combination – and by the network of contracted providers. Health insurers are allowed to selectively contract with health care providers and increasingly do so. As a consequence the number of basic health plans with a limited provider network steadily increased from a single plan in 2008 to 17 plans in 2015, accounting for about 7.5 percent of the insured population [36]. Limited provider plans typically reimburse only part of cost of non-contracted providers [36]. Supplementary insurance does not provide full reimbursement for non-contracted health care providers of care covered by the basic insurance. Citizens with limited provider plans thus need to pay attention to which providers are contracted by their insurer in order to avoid extra payment. The basic idea behind the reform is that insurers negotiate contracts with providers based on price, quantity, and quality of care [5]. To date, however, insurer-provider negotiations are mainly focused on price and volume rather than quality. This is at least partly caused by the absence of reliable quality indicators [12, 37, 38].

Attributes and attribute levels

The first step in a DCE is to carefully select the attributes and levels to be included in the DCE. As the validity of the DCE depends on this process, it needs to be reported clearly as well [39, 40]. We selected four attributes: level of provider choice and associated level of reimbursement, the prime focus of insurers' contracting practices, service benefits provided by the insurer, and monthly premium (Table 1). This selection was based on a literature study (including previous stated-preference studies on the same topic and grey literature), interviews with experts in the field of health insurance (researchers, policy makers, and employees of a health insurance company, n=5) and focus group discussions (FGDs, n=4) with a total number of 23 participants. Each of the four FGDs consisted of at least five

members of the Dutch adult population with mixed characteristics based on age, gender, educational level and health status. Following semi-structured discussions, participants ranked a list of attributes according to the nominal group technique [41]. For this matter, a predefined attribute list (n=9) based on the literature and expert interviews was completed with attributes that were mentioned during the discussions (with a maximum of n=17 attributes to rank).The attributes that were ranked highest were considered for inclusion in the DCE. If not indicated otherwise below, the corresponding attribute levels were chosen based on the health plans available on the market at the time of the study.

As the level of provider choice in the Netherlands is mostly determined by the level of reimbursement for services by non-contracted health care providers, we choose to combine these two aspects into one attribute. At the time of study, health insurers had to cover at least 75 percent of the costs of services, when delivered by health care providers not contracted by the insurer. In 2014 the Dutch government proposed to remove all restrictions on the reimbursement level for non-contracted care in order to reinforce the bargaining position of health insurers vis-a-vis health care providers. This proposal would have allowed insurers to set the reimbursement rate for non-contracted care at zero. However, by the end of 2014 the bill to enact this proposal was rejected in the Senate. We nevertheless also included an attribute level based on zero reimbursement of care by non-contracted providers to assess the potential impact of this dismissed government proposal. Furthermore, based on the expert interviews we selected two different levels for the attribute on prime focus of insurers' contracting practices, namely price and high quality care. A third level, social responsibility, was included because FGD participants frequently mentioned their preference for trustworthy and ethical health insurers – for whom the emphasis is not on financial gains but on a sustainable relationship with the contracted professionals, their enrolees and society. As no general grading system for service benefits exist, we chose positive and negative ratings on consumer websites as attribute levels for the attribute on service benefits provided by the insurer.

Table 1 | Attributes and attribute levels included in the DCE.

Attributes	Attribute levels	Coefficients in regression analysis[1]	Explanation
Provider choice and associated level of reimbursement	50% (0%)	Reference	Percentage of fully reimbursed providers of hospital care and associated percentage of reimbursement of non-contracted providers of hospital care (in brackets)
	80% (75%)	β1	
	100%	β2	
Primary focus of provider contracting	Price of care	Reference	Health insurers contract care providers mainly focussed on either the quality of the care provided, price of the care provided or based on social responsibility
	Quality of care	β3	
	Social responsibility	β4	
Level of service benefits	Negative ratings on consumer websites	Reference	Customer services include the friendliness and ease of contact with the health insurer and the accurate and prompt reimbursement of the claims of consumers
	Positive ratings on consumer websites	β5	
Premium	€70	β6	The monthly nominal premium for a basic health insurance that needs to be paid directly by the consumer to the health insurer.
	€85		
	€100		
	€115		

Notes: [1]The attributes 'provider choice and associated level of reimbursement, 'primary focus of provider contracting', and 'level of service benefits' were dummy coded. Premium was coded as a linear attribute.

DCE design

Each choice set consisted of two basic health plan alternatives, alternative A and alternative B, with systematically varying attribute levels. We chose not to include a 'no basic health insurance' alternative (i.e. opt-out), because basic health insurance is obligatory in the Netherlands. In addition, the DCE did not include an individual-specific status quo alternative because the objective of this study was to estimate trade-offs between health plan characteristics and not the expected uptake of certain basic health plans. As it is not feasible to present a single respondent with all the possible combinations of the included attribute levels, we generated a subset of 30 choice sets (by minimizing the D-error) using Ngene design software (ChoiceMetrics, version 1.1.1.). It was assumed unrealistic to include combinations of the lowest level of 'provider choice and associated level of reimbursement' with the highest 'premium' level and vice versa. The 30 choice sets were blocked into two different questionnaire versions, such that each block has (near) attribute level balance [42]. The D-error of our final design was 0.35 and the S-estimate was 124. The design was optimised for a multinomial logit (MNL) model. Hence, while a loss in efficiency is expected due to the estimation of a latent class model, we expect this loss to be fairly small [43].

Questionnaire

The questionnaire contained two parts. Part one included an introduction to the study topic and the DCE task, a choice set example including an explanation of the terms used, and either one of the blocks of 15 choice sets. In addition, we included two extra choice sets per respondent to test two important axioms underlying DCE methodology [44-46], namely (i) whether or not more attractive attribute levels were preferred over less attractive levels (monotonicity axiom) and (ii) whether or not respondents' preferences were stable (completeness axiom). Thus, each respondent needed to complete 17 choice sets in total. Part two of the questionnaire consisted of questions about the respondent's socio demographic characteristics and previous experiences with health insurance. The questionnaire ended with an open text question where respondents could provide feedback about the questionnaire. We conducted four think aloud interviews to gain insights in how respondents interpreted the questions and answering categories, and whether or not anything was unclear to them. Based on these interviews, we made some minor changes to the wording of the questionnaire. Hereafter, we conducted a formal pilot test with 56 randomly selected internet panel members. Choice data was analysed using a MNL model, and estimates were used to update the priors for the experimental design. Additionally, based on the feedback of respondents on the open text question, some small adjustments to the wording of the questionnaire were made.

Sampling and data collection

A Dutch research company (CG selections) invited 17,000 of their internet panel members to participate in the current study by sending them an email with an URL to the final

questionnaire. Quota sampling was applied, based on age, gender, level of education, and on questionnaire version to ensure an equal distribution of these characteristics over the two questionnaire blocks. Based on the rule of thumb as suggested by Orme [47], we aimed to have at least n=500 respondents to complete the questionnaire. Post hoc analysis showed that this sample size was larger than the minimum required sample size based on the S-estimate. The data was collected during the annual open enrolment period in which Dutch citizens have the opportunity to switch health plans. The questionnaire was online from the 17th of December 2014 to the 6th of January 2015, when the target number of respondents with complete responses was reached.

Statistical analysis

The choice for a particular health plan alternative is modelled as a function of the attributes and their levels in Nlogit (Econometric Software, version 5.0). As one of the aims of our study was to test for preference heterogeneity, both panel latent class models (LCMs) and panel mixed logit models are suitable [48]. The LCM had a better model fit (based on pseudo-R-squared) and is therefore presented in this paper.

LCMs assume the existence of subgroups (called classes) of respondents with homogeneous preferences. Class specific regression coefficients indicate the relative weight respondents of a certain class attach to certain attribute levels. The sign of a coefficient reflects whether the attribute level has a positive or negative effect on health plan preferences. The utility function for basic health plans in the LCM with the best fit was:

Visj|c = β1|c provider choice and associated level of reimbursement (80% (75%)) isj|c+ β2|c provider choice and associated level of reimbursement (100%) isj|c + β3|c contracting (quality) isj|c + β4|c contracting (social responsibility) isj|c + β5|c service level (positive) isj|c + β6|c premium isj|c

The systematic component of utility, V, describes the utility that respondent 'i' belonging to class 'c' has for alternative 's' in choice set 'j'. The attributes 'provider choice and associated level of reimbursement', 'primary focus of provider contracting', and 'level of service benefits' were dummy coded. We included the premium attribute as a linear term. Transformations (such as logarithmic and square root) did not improve model fit.

The researcher pre-specifies the number of classes, based on model fit. Class membership is latent in that the researcher does not determine who belongs to which class a priori. Instead, class membership is expressed in class probabilities that may depend on the respondent's characteristics. Thus, every respondent has a certain probability to belong to one of the latent classes. In addition to the choice model, we fitted a class assignment model to test whether class membership is dependent on age, health status and income. These three

variables were selected based on literature [12-15]. Gross household income per month was measured on a categorical scale. Respondents' answers were subdivided into four groups: low (less than €2000), average (between €2000 and €3000), high (€3000 or more), and does not know, or does not want to tell. Answers for respondents of the latter group were imputed in SPSS (version 22) using single imputation based on regression methods for the DCE analysis. Health status was self-reported: respondents were asked to select whether they suffered from one or more of the 28 chronic disease categories from a predefined list as used by Statistics Netherlands in health research. The final class assignment model was:

$$Vic = \beta 0|c + \beta 1|c \text{ age (30-44 years) } i + \beta 2|c \text{ age (45-59 years) } i + \beta 3|c \text{ age (60+ years) } i + \beta 4|c \text{ chronically ill (yes) } i + \beta 5|c \text{ income (average) } i + \beta 6|c \text{ income (high) } i$$

The reference level for age was the youngest age group (18-29 years), the reference level for health status was having no chronic disease, and the reference level for income was having a low income (less than €2000 per month).

The class specific relative importance of the attributes was calculated by dividing the difference in utility between the highest and lowest level of a single attribute by the sum of the differences of all attributes for that class. The proportions represent the part of someone's preference that is based on that attribute. The higher the percentage the more important the attribute is for health plan choice. Marginal monthly willingness-to-pay (WTP) estimates were calculated by taking the ratio of the change in the attribute level of interest over the negative coefficient of the premium attribute.

RESULTS

Respondent characteristics

The sample (n=533) was representative for the Dutch adult population based on age groups, gender, and educational level (Table 2). Approximately half of the respondents (55.3%) suffered from one or more chronic diseases, which is higher than in the general adult population (35.5%) [49]. Of all respondents, 14.8% did not know what type of health plan they had in 2014 (Table 3). 77.1% of respondents stated to have a supplementary insurance and 57.2% was insured via a group contract, which is lower than in the general population (84% and 70%) [36]. Internet was the most common source of information for those respondents who actively searched for information about health plans and insurance companies because they considered switching, 75.0% stated to use comparison websites, while 69.7% stated to look at the website of the insurance company. Almost half of the sample (45.6%) had switched health plans since the reform in 2006, which is higher compared to the general population (31%) [50]. The latter number includes only people that

switched between insurers, and not those that switched to another health plan of the same insurer, while the former number includes both. The most common reason not to switch to another health plan is 'being satisfied with current plan', followed by 'losing benefits of group contract'.

Table 2 | Demographic characteristics of the study population (n=533).

		Study population	Dutch adult population[1]
Age in years	Mean ± SD	47.0 ± 14.9	
	Range	19-70	
Age group	18-29 years (%)	18.8	19 (17-20)
	30-44 years (%)	25.3	27 (24-29)
	45-59 years (%)	27.6	27 (25-29)
	60+ years (%)	28.3	28 (25-30)
Gender	Female (%)	50.7	51 (46-55)
Educational level[2]	Low (%)	24.0	22 (20-24)
	Average (%)	42.0	42 (38-46)
	High (%)	34.0	34 (31-37)
Gross income per month[3]	Low (%)	34.1	
	Average (%)	19.9	
	High (%)	27.4	
	Does not know, or does not want to tell (%)	18.6	
Chronic diseases	Mean ± SD	1.2 ± 1.5	
	Range	0-9	
	None (%)	44.7	
	1 disease (%)	24.2	
	2 diseases (%)	14.1	
	3 diseases (%)	8.3	
	> 3 diseases (%)	8.8	

Abbreviations: *SD* standard deviation.
Notes: [1]Based on numbers provided by the Center for Information Based Decision Making & Marketing Research (MOA) in cooperation with Statistics Netherlands [51]. Between brackets is the range of proportions that is considered representative for the Dutch adult population for a sample of n=500). [2]Educational level was subdivided into three groups: low (primary education and lower secondary education, average (higher secondary education and intermediate vocational education), and high (tertiary education). [3]Low (less than €2000), average (between €2000 and €3000), high (€3000 or more).

Table 3 | Previous experiences with health insurance (n=533).

		Percentage
Type of health plan in 2014	In-kind benefits	17.3
	Cash benefits	67.9
	Unknown	14.8
Supplementary insurance in 2014	Yes	77.1
	No	21.6
	Unknown	1.3
Group contract in 2014	Yes	57.2
	No	41.1
	Unknown	1.7
Voluntary deductible in 2014	Yes	13.9
	No	84.1
	Unknown	2.1
Time spent annually on considering switching to other health plan	none	32.5
	1 – 29 minutes	17.3
	30 – 59 minutes	17.8
	60 – 179 minutes	21.6
	≥ 180 minutes	10.9
Source of information[1]	Comparison websites	75.0
	Website insurance company	69.7
	Family and friends	28.9
	Consumer federation	20.8
	Call to insurance company	16.1
	Health care provider	6.4
	Other	2.8
Switched health plan since 2006	No	53.1
	Yes	45.6
	Unknown	1.3
Reasons for not switching since 2006[2]	Satisfied with current health plan	79.9
	Losing benefits of group contract	19.1
	Transaction costs	12.0
	Afraid of non-acceptance by health insurer	7.8
	Never thought of it	2.1
	Late with payments	1.8
	Did not know how to switch	0.4
	Other	2.1
To what extent a supplementary insurance affects choices for basic health insurance	Much or very much	60.8
	To some extent	22.5
	Not at all	16.7

Notes: [1]The sum of the percentages is more than 100% as it was possible to give multiple answers. Only those respondents that indicated to spent time on (potential) switching were included to calculate percentages (n=360 (67%)). [2]The sum of the percentages is more than 100% as it was possible to give multiple answers. Percentages were calculated using answers of only those who stated not to have switched health insurance since 2006 (n=283 (53%)).

Analysis of choice data

Only few respondents (7.3%) found the choice sets hard or very hard to answer. Almost all respondents (97.0%) preferred the alternative with better levels to the alternative with worse levels, while 82.4% of respondents had stable preferences, implying that respondents behaved according to the monotonicity and completeness axiom. Almost a quarter of the respondents (22.9%) opted for the alternative with the highest level of provider choice and associated level of reimbursement in every choice set. Few respondents opted for the alternative with the lowest premium (2.4%), positive rating of service levels (1.9%), quality as primary focus of provider contracting (0.8%), or social responsibility as primary focus of provider contracting (0.2%) in every choice set in which it was included.

The choice model with the best fit included four latent classes (Table 4). The model distinguishes a class where freedom of choice dominates (class 1), a class that is service oriented (class 2), a class that is quality minded (class 3), and a cost-conscious class (class 4). The average class probabilities indicate that the majority of respondents was assigned to class 1 (50%), followed by 26% to class 4, and 12% to both class 2 and 3.

The class probability was dependent on age, health status and income (Table 4, class assignment model). Respondents aged 45 or over more often belonged to the class where freedom of choice (and associated reimbursement level) dominates compared to younger respondents who more often belonged to the cost-conscious class ($p<0.01$). Likewise, those who suffered from one or more chronic conditions, and those with a monthly income of €3000 or more had a higher probability to belong to the class where choice dominates than to the cost-conscious class ($p<0.05$). Those with an income of €3000 or more per month were also more likely to belong to the quality minded class compared to those with a lower income, who more often belonged to the cost-conscious class ($p<0.05$).

The coefficients had the expected sign, indicating theoretical validity of the DCE. Respondents across all classes preferred a free choice of care providers to restricted provider choice (and associated reimbursement level). All respondents also preferred selective contracting that was primarily focussed on quality of care over selective contracting focussed on price of care (reference level), whereas contracting focussed on social responsibility was only preferred over the reference level by the service demanders (class 2) and the quality minded responders (class 3). As expected, health plans with positive ratings of service benefits were significantly preferred over plans with negative ratings of service benefits, and lower priced health plans were preferred over higher priced ones.

Table 4 | Final model results panel latent class model (n=533).

		Class 1: choice dominates			Class 2: service demanders			Class 3: quality minded			Class 4: cost-conscious[1]		
Attributes	**Attribute levels**	**β**	**SE**	**RI (rank)**	**β**	**SE**	**RI (rank)**	**β**	**SE**	**RI (rank)**	**β**	**SE**	**RI (rank)**
Provider choice and associated level of reimbursement	50% (0%)(ref)	-	-	0.78 (1)	-	-	0.36 (1)	-	-	0.43 (1)	-	-	0.37 (2)
	80% (75%)	3.60***	0.15		1.82***	0.20		0.48***	0.10		1.68***	0.07	
	100%	6.55***	0.23		3.11***	0.34		1.78***	0.19		2.70***	0.12	
Primary focus of provider contracting	Price(ref)	-	-	0.02 (4)	-	-	0.13 (4)	-	-	0.33 (2)	-	-	0.03 (4)
	Quality	0.20**	0.08		1.14***	0.16		1.42***	0.10		0.22***	0.06	
	SR	-0.09	0.07		0.31**	0.12		0.22***	0.08		0.11*	0.06	
Level of service benefits	Negative (ref)	-	-	0.11 (2)	-	-	0.35 (2)	-	-	0.17 (3)	-	-	0.09 (3)
	Positive	0.89***	0.07		2.94***	0.16		0.69***	0.07		0.66***	0.04	
Premium per month	Per 10 €	-0.17***	0.03	0.09 (3)	-0.29***	0.05	0.15 (3)	-0.06**	0.03	0.07 (4)	-0.81***	0.02	0.51 (1)
Class assignment model		**Par.**	**SE**		**Par.**	**SE**		**Par.**	**SE**		**Par.**	**SE**	
Constant		-0.64**	0.28		-1.71***	0.46		-1.82***	0.50		-	-	
Age categories	18-29 (ref)	-	-		-	-		-	-		-	-	
	30-44 years	0.58*	0.34		0.55	0.52		0.45	0.59		-	-	
	45-59 years	1.04***	0.36		0.77	0.53		0.94	0.60		-	-	
	60+ years	1.25***	0.37		0.90	0.55		1.06*	0.60		-	-	
Chronically ill	No (ref)	-	-		-	-		-	-		-	-	
	Yes	0.50**	0.25		0.26	0.36		-0.19	0.38		-	-	
Income[2]	Low (ref)	-	-		-	-		-	-		-	-	
	Average	0.36	0.30		0.43	0.44		0.48	0.52		-	-	
	High	0.61**	0.29		0.54	0.42		1.11**	0.45		-	-	
Average class probability		0.50			0.12			0.12			0.26		
Model fit													
Number of observations		7995											
AIC		0.92											
Pseudo-R-squared[3]		0.35											
Log likelihood		-3624											

Abbreviations: *ref* reference; *SR.* social responsibility; *Par.* parameter; *SE* standard error; *RI* relative importance.
*, **, *** denote significance at the 0.10, 0.05 and 0.01 level, respectively.
Notes: [1]Class 4 does not have parameters in the class assignment model as the parameters of class 1 to 3 are relative to class 4. [2] Answers for respondents of the 'does not know, or does not want to tell' group were imputed in SPSS using single imputation based on regression methods. The new distribution of income was: low income 40.3%, average income 25.1%, and high income 34.5%. [3]Pseudo-R-squared is not the same as R-squared that is used in linear regression. A pseudo-R-squared of 0.3-0.4 is equivalent to a R-squared between 0.6 and 0.8 [21].

Table 5 | Class specific marginal monthly willingness-to-pay estimates in euros.

		Class 1: choice dominates		Class 2: service demanders		Class 3: quality minded		Class 4: cost-conscious	
Attributes	**Attribute levels**	**WTP**	**95% CI[1]**	**WTP**	**95% CI[1]**	**WTP**	**95% CI[1]**	**WTP**	**95% CI[1]**
Provider choice and associated level of reimbursement	50% (0%)→ 100%	396	270;522	108	79;136	279	68;490	33	31;35
	80% (75%)→ 100%	178	118;239	45	26;63	204	41;367	13	11;15
Primary focus of provider contracting	Price → quality	12	2;23	39	25;53	222	32;412	3	1;4
	Price → SR	-6	-14;3	11	1;20	34	-4;71	1	0;3
Level of service benefits	Negative → positive	54	34;74	102	72;132	108	22;195	8	7;9

Abbreviations: *CI.* confidence interval; *SR.* social responsibility.
Notes: [1]95% confidence intervals were calculated using the Delta Method [52]

The relative value attached to health plan attributes differed between the classes (Table 4). Provider choice and associated reimbursement level was, compared to the other three attributes, by far the most important attribute in class 1. In class 2, provider choice and associated reimbursement level was also the most important attribute; however, this was followed closely by the service benefits provided by the insurer. The primary focus of provider contracting was, besides the level of provider choice and associated reimbursement level, the most important attribute for respondents of class 3. For respondents of class 4, monthly premium was most important, followed by level provider choice and associated reimbursement level.

The class specific WTP estimates are all statistically significant, except the estimates for a change in primary focus of provider contracting from price to social responsibility in class 1 and 3 (Table 5). Estimates differ importantly between the classes, except for class 2 and 3 where confidence intervals overlap for all attributes. The high WTP estimates in class 1 are in direct contrast with the lower estimates in class 4. These high WTP estimates in class 1 indicate that those respondents will never choose a health plan with restricted provider choice and lower reimbursement levels. Except for class 3, WTP estimates for a change in the primary focus of provider contracting from price to quality are low, indicating that consumers do not value the focus of insurers on quality of care highly. The wide CIs in class 3 indicate heterogeneous preferences within this class, while the narrow CIs in class 4 indicate homogenous preferences.

DISCUSSION

The level of provider choice and the associated reimbursement level was by far the most decisive health plan characteristic for respondents aged 45 or over, those with one or more chronic conditions, and those with a higher income. Although this characteristic was also important for those young, healthy, and with a low or average income, monthly premium was their most important determinant of choice. Health insurers with a focus on contracting high quality providers were preferred rather than lower-priced providers, although the WTP estimates (€12, €39, €222 and €3 for class 1, 2, 3 and 4 respectively) were low compared to the other WTP estimates of a given class, except for class 3. A minority of the respondents preferred insurers with a focus on social responsibility rather than price when contracting providers.

Interpretation of the results

A possible explanation for the dominance of the provider choice attribute in the current study might be that respondents prefer to choose health care providers regardless of financial considerations. In addition, the dominance might partly be related to an objection

to the government proposal allowing insurers to set a lower level of reimbursement of care by non-contracted providers. At the time of data collection, this proposal was heavily disputed by the medical profession and widely covered by the media. Restricting provider choice was also one of the reasons for the backlash against Managed Care Organizations in the US in the 1990s [53, 54]. A DCE study in the US setting also found that the physician network was the most important attribute for health plan choice, followed by prescription coverage and costs per visit [24]. Earlier Dutch DCEs found a negative WTP of €76 per year for physician choice from a predefined network, and of €137 for choice from a list established by the health insurer, compared to physician choice based on the gatekeeper model (i.e. a model in which general practitioners decide on referral of patients to medical specialists). From these studies it follows that the WTP for free physician choice was positive (€79) [29, 30]. Our results differ from the most recent DCE study that found no effect of the level of provider choice on insurance preferences [33]. Bergrath and colleagues suggested that the insignificance had to do with consumers being not fully aware of all the aspects that play a role in a health care system with managed competition. Our study confirms earlier DCE findings of preference heterogeneity with respect to restrictions on provider choice and premium [29, 30, 33]. Our results indicate that consumers do not value the focus of insurers on quality of care highly. This might indicate that consumers do not expect insurers to play an important role as purchasing agents of high quality care. Instead, consumers seem to rely on their general practitioner to select which are the best care providers to go to [55].

Strengths and limitations

The paper contributes to the existing literature in three different ways. First, our study provides insights in (unobserved) consumer preferences about health plan characteristics after almost ten years of experience with managed competition in the Netherlands. Second, relevant preferences for plan characteristics were identified by a combination of extensive literature search, expert interviews and focus groups discussions. Third, in contrast to the most recent Dutch DCE study in the field of health plans, our study is based on a large representative sample of the national adult population. This study is also subject to some limitations. First, we specified only one aspect of provider choice (i.e. reimbursement level) whereas other aspects, such as the location of the (non-)contracted providers, may be relevant as well. More specifically, respondents might not care about the level of provider choice (and associated reimbursement level) when the health plan contracts all nearby hospitals. In that case, other health plan characteristics, such as premium, might become more important for their decision. Second, because only a limited number of attributes and attribute levels can be included in a DCE we only selected the four most important attributes based on the qualitative study. Given that for the majority of respondents characteristics of supplementary health insurance affect their choice for basic health insurance much or very much (60.8%, Table 3) it is unfortunate that because of the large heterogeneity of supplementary insurance products we were not able to take the role of

a supplementary insurance in health plan choice explicitly into account. Third, as many different implementations of basic health plans exist in the Netherlands, it was not possible to define one specific status quo alternative that would be applicable to all respondents. In addition, we chose not to include an individual-specific status quo alternative. This absence of a status quo in this context prevents to draw conclusions on uptake of different health plans based on stated preferences of consumers. Although calculation of uptake rates is a very useful application of DCE data, we feel that the information on trade-offs that consumers make between different features of health plans is useful in itself, both for policy makers and health insurance companies.

Policy implications

Despite these limitations, this study provides useful evidence on consumer preferences, and the heterogeneity herein, that can inform policy decisions on the design and regulation of the insurance market. Although consumers strongly prefer free choice of provider, respondents still trade off free provider choice against premium. Given the high WTP estimates in class 1 however, it is likely that half of the respondents will never opt for restricted health plans. Apparently, health insurers will not be able to steer these respondents to preferred providers. Conversely, respondents belonging to class 4 (class probability: 26%) are willing to accept limited provider choice for a monthly premium discount of €13 (95% CI: €11; €15), as well as some respondents from class 2 (class probability: 12%) given that the lower level of the 95% CI is €26. These estimates are larger than the €7 difference in monthly premiums between restricted and non-restricted health plans in the year of the study, 2014 [36]. Insurers lowering the premium for choice restricted health plans might therefore attract a selective group of new clients, namely the young, healthy and less wealthy consumers. If insurers are able to provide objective quality information of the contracted providers that is easy to understand, consumers might make more well-considered insurance choices that do include quality of contracted care. As a result, consumers might be more willing to accept less freedom of provider choice. Recently, the Dutch minister of Health advised the insurers to be more transparent about their contracting practices [56].

CONCLUSION

We conclude that it would be very unlikely for half of the sample to opt for health plans with restricted provider choice. However, a premium discount up to €15/month by restricted health plans might motivate especially younger, healthier, and less wealthy consumers to choose these plans.

REFERENCES

1. Enthoven AC, van de Ven WP. Going Dutch--managed-competition health insurance in the Netherlands. New Engl J Med. 2007;357:2421-3.
2. Health Insurance in the Netherlands. Den Haag: Ministry of Health, Welfare and Sports, 2011. Available from: http://www.government.nl/files/documents-and-publications/leaflets/2012/09/26/health-insurance-in-the-netherlands/health-insurance-in-the-netherlands.pdf.
3. OECD Reviews of Health Systems: Switzerland 2011. Paris: Organisation for Economic Co-operation and Development (OECD)/ World Health Organization (WHO), 2011. Available from: http://www.oecd.org/els/health-systems/oecdreviewsofhealthsystems-switzerland.htm.
4. Leu RE, Rutten FH, Brouwer W, et al. The Swiss and Dutch health insurance systems: universal coverage and regulated competitive insurance markets. Commonwealth Fund; 2009. Available from http://www.amcp.org/WorkArea/DownloadAsset.aspx?id=13049.
5. Schafer W, Kroneman M, Boerma W, et al. The Netherlands: health system review. Health systems in transition. 2010;12:v-xxvii, 1-228.
6. Maarse H, Meulen RT. Consumer choice in Dutch health insurance after reform. Health care Anal. 2006;14:37-49.
7. Abaluck J, Gruber J. Choice Inconsistencies Among the Elderly: Evidence from Plan Choice in the Medicare Part D Program. Am Econ Rev. 2011;101:1180-210.
8. Heiss F, Leive A, McFadden D, et al. Plan selection in Medicare Part D: evidence from administrative data. J Health Econ. 2013;32:1325-44.
9. Kling JR, Mullainathan S, Shafir E, et al. Comparison friction: experimental evidence from medicare drug plans. Q J Econ. 2012;127:199-235.
10. Handel BR, Kolstad JT. Health Insurance for" Humans": Information Frictions, Plan Choice, and Consumer Welfare. Am Econ Rev. 2015;105:2449-500.
11. Handel BR. Adverse selection and inertia in health insurance markets: When nudging hurts. Am Econ Rev. 2013;103:2643-82.
12. Brabers AE, Reitsma-van Rooijen M, de Jong JD. The Dutch health insurance system: mostly competition on price rather than quality of care. Eurohealth. 2012;18:30.
13. Boonen LH, Laske-Aldershof T, Schut FT. Switching health insurers: the role of price, quality and consumer information search. Eur J Health Econ. 2016;17:339-53.
14. Reitsma-van Rooijen M, Brabers AEM, de Jong JD. [Free choice of healthcare providers is of importance when choosing health plans]. Vrije keuze van zorgaanbieders van belang bij het kiezen van een polis. Utrecht: Netherlands Institute for Health Services Research (NIVEL), 2015. Available from: http://www.nivel.nl/sites/default/files/bestanden/Factsheet-wisselen-van-zorgverzekeraar-2015.pdf. *In Dutch.*
15. Gruber J, McKnight R. Controlling health care costs through limited network insurance plans: Evidence from Massachusetts state employees. Cambridge: National Bureau of Economic Research, 2014. Available from: http://www.nber.org/papers/w20462.
16. Clark MD, Determann D, Petrou S, et al. Discrete Choice Experiments in Health Economics: A Review of the Literature. Pharmacoeconomics. 2014;32:883-902.
17. Luce D, Tukey J. Simultaneous conjoint measurement: a new type of fundamental measurement. J Math Psychol. 1964;1:1-27.
18. McFadden D. Conditional logit analysis of qualitative choice behavior. In: Zarembka P, editor. Frontiers of Economics. New York: Academic Press; 1974. p. 105-42.
19. Manski CF. Structure of Random Utility Models. Theor Decis. 1977;8:229-54.
20. Lancsar E, Louviere J. Conducting discrete choice experiments to inform healthcare decision making: a user's guide. Pharmacoeconomics. 2008;26:661-77.
21. Hensher DA, Rose JM, Greene WH. Applied choice analysis: a primer. New York: Cambridge University Press; 2005.

22. Abiiro GA, Torbica A, Kwalamasa K, et al. Eliciting community preferences for complementary micro health insurance: a discrete choice experiment in rural Malawi. Soc Sci Med. 2014;120:160-8.
23. Trujillo AJ, Ruiz F, Bridges JF, et al. Understanding consumer preferences in the context of managed competition: evidence from a choice experiment in Colombia. Appl Health Econ Health Pol. 2012;10:99-111.
24. Gates R, McDaniel C, Braunsberger K. Modeling consumer health plan choice behavior to improve customer value and health plan market share. J Bus Res. 2000;48:247-57.
25. Kesternich I, Heiss F, McFadden D, et al. Suit the action to the word, the word to the action: Hypothetical choices and real decisions in Medicare Part D. J Health Econ. 2013;32:1313-24.
26. Becker K, Zweifel P. Age and choice in health insurance: evidence from a discrete choice experiment. Patient. 2008;1:27-40.
27. Kerssens JJ, Groenewegen PP. Consumer preferences in social health insurance. Eur J Health Econ. 2005;6:8-15.
28. Van den Berg B, Van Dommelen P, Stam P, et al. Preferences and choices for care and health insurance. Soc Sci Med. 2008;66:2448-59.
29. Leukert-Becker K, Zweifel P. Preferences for health insurance in Germany and the Netherlands - a tale of two countries. Health Econ Rev. 2014;4:22.
30. MacNeil Vroomen J, Zweifel P. Preferences for health insurance and health status: does it matter whether you are Dutch or German? Eur J Health Econ. 2011;12:87-95.
31. Boonen LH, Schut FT, Donkers B, et al. Which preferred providers are really preferred? Effectiveness of insurers' channeling incentives on pharmacy choice. Int J Health Care Finance Econ. 2009;9:347-66.
32. Boonen LH, Donkers B, Schut FT. Channeling consumers to preferred providers and the impact of status quo bias: does type of provider matter? Health Serv Res. 2011;46:510-30.
33. Bergrath E, Pavlova M, Groot W. Attracting Health Insurance Buyers through Selective Contracting: Results of a Discrete-Choice Experiment among Users of Hospital Services in the Netherlands. Risks. 2014;2:146-70.
34. Van de Ven WP, Schut FT. Universal mandatory health insurance in the Netherlands: a model for the United States? Health Aff (Millwood). 2008;27:771-81.
35. Shmueli A, Stam P, Wasem J, et al. Managed care in four managed competition OECD health systems. Health Policy. 2015;119:860-73.
36. [Marketscan Health Insurance Market 2015 - Part A]. Marktscan Zorgverzekeringsmarkt 2015 - Deel A. Utrecht: The Dutch Healthcare Authority (NZa), 2015. Available from: https://www.nza.nl/1048076/1048181/Marktscan_Zorgverzekeringsmarkt_2015__deel_A_en_brief.pdf. *In Dutch.*
37. Van Ginneken E, Swartz K. Implementing insurance exchanges--lessons from Europe. N Engl J Med. 2012;367:691-3.
38. [Monitor and Policy letter on Purchasing of Care]. Monitor en beleidsbrief Zorginkoop. Utrecht: The Dutch Healthcare Authority (NZa), 2014. Available from: https://www.nza.nl/104107/105773/953131/Monitor_Zorginkoop_en_beleidsbrief.pdf. *In Dutch.*
39. Abiiro GA, Leppert G, Mbera GB, et al. Developing attributes and attribute-levels for a discrete choice experiment on micro health insurance in rural Malawi. BMC Health Serv Res. 2014;14:235.
40. Coast J, Horrocks S. Developing attributes and levels for discrete choice experiments using qualitative methods. J Health Serv Res Policy. 2007;12:25-30.
41. Hiligsmann M, van Durme C, Geusens P, et al. Nominal group technique to select attributes for discrete choice experiments: an example for drug treatment choice in osteoporosis. Patient Pref Adherence. 2013;7:133-9.
42. Reed Johnson F, Lancsar E, Marshall D, et al. Constructing experimental designs for discrete-choice experiments: report of the ISPOR Conjoint Analysis Experimental Design Good Research Practices Task Force. Value Health. 2013;16:3-13.
43. Bliemer MC, Rose JM. Construction of experimental designs for mixed logit models allowing for correlation across choice observations. Transport Res B-Meth. 2010;44:720-34.

44. Cairns J, van der Pol M. Repeated follow-up as a method for reducing non-trading behaviour in discrete choice experiments. Soc Sci Med. 2004;58:2211-8.
45. Lancsar E, Louviere J. Deleting 'irrational' responses from discrete choice experiments: a case of investigating or imposing preferences? Health Econ. 2006;15:797-811.
46. Lloyd AJ. Threats to the estimation of benefit: are preference elicitation methods accurate? Health Econ. 2003;12:393-402.
47. Orme BK. Getting started with Conjoint Analysis: Strategies for Product Design and Pricing Research. Madison, WI: Reserach Publishers LLC; 2006.
48. Hole AR. Modelling heterogeneity in patients' preferences for the attributes of a general practitioner appointment. J Health Econ. 2008;27:1078-94.
49. Statistics Netherlands (CBS). [Health and Healthcare use; personal characteristics]. Gezondheid en zorggebruik; persoonskenmerken. Available from: http://statline.cbs.nl/Statweb/publication/?DM=SLNL&PA=83005ned&D1=0-2,6-29,46-51&D2=6-13,16&D3=a&D4=l&HDR=G1&STB=G2,G3,T&VW=T. [Last accessed 09 06 2016]. *In Dutch.*
50. ['Zorgthermometer': overview of the insured in 2015]. Zorgthermometer: Verzekerden in beeld 2015. Zeist: Executive Organization of Health Insurance Companies in the Netherlands (Vektis), 2015. Available from: http://www.vektis.nl/downloads/Publicaties/2015/Zorgthermometer%20nr14/index.html. *In Dutch.*
51. MOA - Center for Information Based Decision Making & Marketing Research. Available from: http://www.moaweb.nl/english. [Last accessed 02 10 2015].
52. Hole AR. A comparison of approaches to estimating confidence intervals for willingness to pay measures. Health Econ. 2007;16:827-40.
53. Robinson JC. The end of managed care. JAMA. 2001;285:2622-8.
54. Mechanic D. The rise and fall of managed care. J Health Soc Behav. 2004;45 Suppl:76-86.
55. Rademakers J, Nijman J, Brabers AE, et al. The relative effect of health literacy and patient activation on provider choice in the Netherlands. Health Policy. 2014;114:200-6.
56. [Letter from the Minister to the House of Representatives 'Transparancy health insurers']. Kamerbrief 'Transparantie zorgverzekeraars' van Minister Schippers aan de Tweede Kamer. Ministry of Health, Welfare and Sports, 2015. Available from: http://www.rijksoverheid.nl/documenten-en-publicaties/kamerstukken/2015/06/30/kamerbrief-over-transparantie-zorgverzekeraars.html. [Last accessed 24 08 2015]. *In Dutch.*

7

Personal Health Records in the Netherlands: potential user preferences quantified by a discrete choice experiment

Domino Determann
Mattijs S. Lambooij
Dorte Gyrd-Hansen
Esther W. de Bekker-Grob
Ewout W. Steyerberg
Marcel Heldoorn
Line Bjørnskov Pedersen
G. Ardine de Wit

Submitted

ABSTRACT

Objective: The aim of this study was to investigate personal health record (PHR) preferences of potential users.

Materials and Methods: We performed a discrete choice experiment (DCE), which consisted of 12 choice scenarios, each comprising 2 hypothetical PHR alternatives and an opt-out. The PHR alternatives differed based on five characteristics. The survey was administered to Internet panel members of the Dutch Federation of Patients and Consumer Organizations (NPCF). We used latent class models to analyse the data.

Results: 1,443 potential PHR users completed the DCE. We identified three latent classes: “refusers” (class probability 43%), “eager adopters” (37%) and “reluctant adopters” (20%). The predicted uptake for the reluctant adopters ranged from 4% in case of a PHR with the worst attribute levels to 68% in the best case. Those with one or more chronic diseases were significantly more likely to belong to the eager adopters’ class. The facilitating party for storage of PHR was the most decisive aspect for the eager and the reluctant adopters, while the costs were most decisive for the refusers. Across all classes, healthcare providers, and independent organisations were the most preferred facilitators of data storage.

Discussion and Conclusion: More than one third of potential PHR users indicate great interest in a PHR, irrespective of its characteristics. Policymakers who aim to expand the use of PHRs will be most successful when healthcare providers and health facilities, or independent organizations facilitate the storage of PHR data, while refraining to include market parties in storage of the data.

BACKGROUND AND SIGNIFICANCE

Personal health records (PHRs) can be defined as: *"An electronic application through which individuals can access, manage and share their health information, and that of others for whom they are authorized, in a private, secure, and confidential environment"* [1]. In contrast to the clinician's medical record, PHRs are managed by patients [2]. PHRs can have varying characteristics, for example PHRs can offer secure patient-provider communication, and hold various sources of static and dynamic information for patients [3]. Furthermore, both the provider and the patient may be able to add data [4] either self-entered or through "wearables". Additional convenience features include online appointment scheduling and requesting medication refills. Large scale implementation of PHRs [3, 5] is expected to yield cost reductions [6, 7], increase quality of care [8] and increase efficiency of care [3]. They are seen as a tool to empower patients and give patients control over their healthcare process [9]. As such, PHRs can serve as a clinical information system in patient-centred care, and support well-informed, activated and empowered patients [6].

PHRs are generating an increased interest and are high on the agenda of policy makers in the Netherlands [10]. In contrast to other countries, such as Australia in which a national PHR was implemented in 2012 [11], there is currently no national PHR initiative in the Netherlands. However, a number of PHR platforms are available, e.g. Microsoft's Health Vault and Patient1 [12]. Comparable platforms in other countries include Dossia HMS in the USA, and Patients Know Best in the UK [13]. Involving potential users in the development, testing, and implementation of PHRs could provide a base for further implementation of PHRs [14]. Furthermore, insights into the value that citizens place on various PHR characteristics can help developers to improve their products and advice policy makers on furthering conditions for the use of PHRs [2, 15].

To our knowledge, there is little quantitative information available on the number of users of PHRs in developed countries. However, sources report that the current uptake of PHRs is limited [16, 17]. A recent study among members of a panel of patients and consumers in the Netherlands reports that 9% of the people have a PHR [18]. It has been argued that PHRs that are currently available on the market are not based on patients' needs [19]. The limited uptake may be an important reason for the scarce evidence of PHR advantages [2, 20-24]. If a significant diffusion of PHRs is to take place, it is essential that potential users envisage sufficient added value of PHRs relative to the status quo [25]. Individuals with poorer digital literacy are less inclined to adopt PHRs [14, 25, 26]. Moreover, chronically ill people, frequent healthcare users, and those with a role as caregiver are more likely to use PHRs [27-31]. However, the design and usability of PHRs will affect the eventual uptake [31], also for these groups. A barrier for PHR use expressed in many studies is the concern about security and privacy of the data in the PHR [14, 25, 26, 29, 32, 33]. Another barrier is the

lack of standardisation of the formats in operational systems [34], i.e. data must be entered manually into most PHRs.

OBJECTIVE

In this study, we aim to investigate potential users' preferences for PHR characteristics. Individuals who face the decision to opt for a PHR trade off the negatively valued aspects with positively valued aspects of PHRs. This study aims to mimic this trading behaviour by imitating the real choice situation, adopting or not adopting a PHR, by means of the discrete choice experiment (DCE) methodology [35]. A DCE is a survey-based stated-preference method that has been increasingly used in healthcare to address a wide range of policy questions over the past years [36]. To date, no DCEs have been conducted to explore PHR preferences. The research questions of this paper are: 1) Can we identify subgroups of potential users with different PHR preferences across groups? 2) What PHR characteristics are most important within subgroups?, and 3) What is the potential uptake of PHRs?

MATERIALS AND METHODS

Discrete choice experiments

DCEs originate from mathematical psychology [37] and have a strong theoretical foundation in random utility theory (RUT, [38, 39]) and Lancaster's theory of value [40]. In a typical DCE survey, respondents are presented with a series of hypothetical scenarios (called choice sets) in which they are asked to choose between two or more alternatives that are distinguished from each other by systematically varying characteristics (called attributes) [41-44]. RUT assumes that respondents will choose the alternative within a choice set that yields the highest utility (benefit or satisfaction). Based on respondents' choices, preferences can be elicited, willingness-to-pay (WTP) estimates can be calculated (if a cost attribute is included), and uptake rates can be predicted.

Experimental DCE design

We selected five attributes with two to six levels each (Table 1) based on literature search, expert interviews (policy makers, researchers; N=5), and focus group discussions with the general population (N=4, 25 participants in total). Each of the four focus group discussions consisted of at least six members of the Dutch adult population with mixed characteristics based on age, gender, educational level and health status. Following semi-structured discussions, participants ranked a list of attributes according to the nominal group technique [45]. For this matter, a predefined attribute list (n=17) based on the literature and expert interviews was completed with additional attributes that were mentioned during the focus

group discussions (with a maximum of n=23 attributes to rank). The attributes that were ranked highest were considered for inclusion in the DCE.

Table 1 | Attributes and corresponding levels included in the DCE.

Attributes	Short name	Attribute levels
Facilitating party for storage of PHR data	Facilitator	Commercial company
		Independent organization or platform
		Government
		Healthcare providers or health facility
Level of connectivity	Connectivity	Standalone
		Tethered with the system of your GP
		Tethered with the system of your hospital
		Interconnected
Use of anonymized data by third parties	Use of data	No, never
		Yes, after permission
Option to upload the person's own data into PHR	Data adding	No
		Yes
PHR costs per capita per year as a proxy of the price that persons would pay for a PHR[1]	Costs	€0, €15, €30, €50, €70, €95

Abbreviations: *PHR* personal health record; *GP* general practitioner.
Notes: [1]We included the following costs levels in the pilot survey based on earlier research by the NPCF [18]: €0, €2, €5, €15, €30 and €50. The results of the pilot study showed the need to expand the range of cost levels by increasing the highest level, and to broaden the intervals between the levels.

A choice set consisted of two unlabelled hypothetical PHR alternatives; PHR A and PHR B, with systematically varying attribute levels, and an opt-out alternative (i.e. no PHR). This opt-out was necessary since, as in real life, respondents are not obliged to have a PHR. Additional file 1 shows an example of a choice set. By minimizing the D-error, a Bayesian efficient design based on best guess priors was generated using Ngene design software (version 1.1.1.). These priors (small positive and negative values) were used to increase the efficiency of the design by avoiding dominant choice sets. We generated a subset of 12 choice sets to ensure level balance, as well as enough degrees of freedom to estimate a main effects only model (assuming that the cost attribute is non-categorical). For more information on these more technical elements of a DCE, see e.g. Reed Johnson et al. [46].

Survey instrument

An inclusion criterion for the current study was that respondents currently did not have a PHR. The survey started with an introduction to the study and the topic (Additional file 1). Hereafter, the included attributes and their corresponding levels as well as the DCE task were explained to respondents, followed by a clearly explained choice set example. Subsequently, respondents completed the 12 choice sets. An example of a choice set is

shown in Additional file 1. The survey concluded with additional questions on respondents' intention to use a PHR, questions on sociodemographic characteristics, a question on whether or not respondents had previously experienced a medical error, and questions on digital literacy (internet use, type of internet use and risky digital behaviour). The survey was pilot tested using think aloud interviews (N=4) to test for respondent understanding and to improve the wording of the survey. Next, a formal pilot test (N=51, from the same Internet panel as used for the final data collection, see below) was conducted to test randomisation and to improve the wording of the questionnaire. Pilot data was analysed using multinomial logit models and estimates were used as priors for the final DCE design.

Data collection

The survey was administered to Internet panel members of the Dutch Federation of Patient and Consumer Organizations (NPCF) in December 2015 and January 2016. All panel members (N=22,841) received an email with the URL to the online survey and were asked to complete it on a voluntary basis. Data-collection was terminated once the number of respondents completing the survey had decreased to a few respondents per day. Because of practical limitations at the NPCF, no reminders were sent out. Formal testing of the study protocol by a Medical Ethics Committee was not necessary, as a questionnaire study amongst volunteers of an Internet panel does not fall within the scope of the Dutch Act on Medical Research Involving Human Subjects.

Discrete choice data analysis

The choices respondents made in the choice sets were used to estimate the impact of the attributes (independent variables) on the respondents' choices for PHR or opting-out (dependent variable). A significant independent variable indicates that the attribute level has a significant impact on PHR preferences and the sign of the coefficient reflects whether this impact has a positive or negative effect. Several types of discrete choice models can be estimated. We have chosen to perform a latent class analysis, since this is a model that can take both preference heterogeneity and the panel nature of the data (i.e. dependencies between choice observations by a single respondent) into account [47]. In addition, it is a closed form model (i.e. does not rely on complex simulations) [47]. A latent class analysis assumes the existence of subgroups (called classes) of respondents that differ with respect to preferences. The researcher decides on the number of classes based on the model fit (AIC, BIC, pseudo-R^2) and sound interpretation of classes. Class membership is latent in that the researcher does not determine who belongs to which class a priori. Instead, class membership is expressed as class probabilities that may depend on the respondent's characteristics. We hereto fitted a class probability model in addition to the choice model where we tested the presence of individual-level drivers (age, health status, previous experience of a medical error) and individual-level barriers for the use of PHR (health literacy and digital literacy (internet use, type of internet use and risky digital behaviour))

as predictors of the classes. These characteristics were selected based on literature [14, 25-31]. Choice data was analysed in Nlogit version 4.0 (Econometric Software).

The class specific relative importance of the attributes was calculated by dividing the difference in utility between the highest and lowest level of a single attribute by the sum of the differences in utility of all attributes for that class. The higher the proportion the more important the attribute is for PHR choice. We calculated annual WTP estimates by taking the ratio of the attribute level of interest over the negative coefficient of the cost attribute. We calculated the class-specific uptake of PHRs by taking the exponential of the total utility for a particular PHR divided by the exponential of the sum of utilities for the particular PHR and the opt-out (no PHR). In addition, we calculated the average uptake (weighted average, based on class probabilities).

RESULTS

Descriptive statistics

1,443 panel members fulfilled the inclusion criteria. Table 2 shows that respondents were on average 61 years of age, and half of the sample was highly educated (50%). The majority of respondents suffered from one or more chronic diseases (77%). Compared to the general population, our sample is older, more highly educated and more often chronically ill. About one third of respondents had at least once experienced a medical error (e.g. receiving a wrong drug). The stated intention to use a PHR within two years from now was high for a quarter of the sample (25%).

Discrete choice data

The improvement in model fit (AIC, BIC, pseudo-R^2) was very limited for a model with more than three classes. The final latent class model therefore included three latent classes (Table 3). The average class probabilities were: 43% for class 1 (we refer to this class as “refusers”), followed by 37% for class 3 (referred to as “eager adopters”), and 20% for class 2 (referred to as “reluctant adopters”). Tests for class probability showed that those with one or more chronic diseases had a significantly higher probability of belonging to class 3 compared to class 1 (p-value 0.03) and class 2 (p-value <0.01). The other individual-level drivers and barriers were not significant class probability predictors.

Table 2 | Descriptive statistics.

		Sample (N=1,443)	Census data[1]
		Mean (SD)	**Mean**
Age		61 (11)	41.3
		N (%)	**%**
Age groups in years	18-40	57 (4.3)	33.7
	40-65	704 (53)	44.0
	65-80	533 (40)	16.9
	80 or older	44 (3.3)	5.5
Gender	Female	698 (51)	50.5
	Male	671 (49)	49.5
Education[2]	Low	254 (18)	32.9
	Average	443 (32)	39.3
	High	706 (50)	27.8
Health status	Healthy	325 (23)	69.9
	One or more chronic diseases	1118 (77)	30.1
Subjective health literacy[3]	Low	29 (2.0)	-
	Adequate	1395 (98)	-
Experienced a medical error	No	926 (65)	-
	Yes	500 (35)	-
Digital literacy		**N (%)**	
Internet use	Easy	1187 (83)	-
	Not easy, not hard	216 (15)	-
	Hard	23 (1.6)	-
Type of Internet use	Mainly for fun	765 (53)	-
	Mainly for other purposes	678 (47)	-
		Mean (SD)	
Risky digital behaviour[4]		10.9 (4.4)	-
PHR		**N (%)**	
Stated intention to use a PHR within 2 years from now	Low (0-4)	727 (51)	-
	Average (5 or 6)	346 (24)	-
	High (7-10)	364 (25)	-

Abbreviations: *SD* standard deviation; *PHR* personal health record
Notes: [1]Based on numbers provided by Statistics Netherlands, 2015 [48] [2]Educational level was categorized into three groups: low (primary education and lower secondary education), average (higher secondary education and intermediate vocational education), and high (tertiary education). [3]Subjective health literacy was measured based on the validated Dutch questions of the Set of Brief Screening Questions (SBSQ-D) of Chew [49]. [4]Risky digital behaviour was measured by means of four statements where respondents had to mark on a scale of 1-7 how (un)likely this behaviour was for them (maximum score is thus 28). We hereto adapted the statements included in the domain-specific risk-taking (DOSPERT) scale for adult populations [50] to fit our research question.

Table 3 | Latent class model results (N=1,443).

		Class 1: the refusers				Class 2: the reluctant adopters				Class 3: the eager adopters			
Short name of attribute[1]	**Attribute level**	**β**	**SE**	**P-value**	**RI**	**β**	**SE**	**P-value**	**RI**	**β**	**SE**	**P-value**	**RI**
Constant	No PHR	3.57	0.13	<0.01	N/A	0.51	0.03	<0.01	N/A	-4.13	0.09	<0.01	N/A
Facilitator	Commercial company	-0.50	0.17	<0.01	0.17 (2)	-0.83	0.04	<0.01	0.32 (1)	-0.49	0.03	<0.01	0.31 (1)
	Independent organization or platform	0.38	0.14	<0.01		0.44	0.04	<0.01		0.27	0.03	<0.01	
	Government	0.02	0.17	0.91		-0.06	0.04	0.10		-0.10	0.03	<0.01	
	Care providers or health facility	0.10	0.15	0.52		0.45	0.03	<0.01		0.32	0.03	<0.01	
Connectivity	Standalone	0.10	0.15	0.52	0.10 (4)	-0.28	0.04	<0.01	0.13 (4)	-0.31	0.02	<0.01	0.20 (3)
	Tethered with GP system	-0.14	0.17	0.39		0.24	0.04	<0.01		0.20	0.03	<0.01	
	Tethered with hospital system	0.28	0.14	0.04		-0.15	0.04	<0.01		0.09	0.03	<0.01	
	Interconnected	-0.24	0.17	0.16		0.19	0.04	<0.01		0.02	0.02	0.36	
Use of data	No, never	-0.12	0.10	0.24	0.05 (5)	-0.24	0.02	<0.01	0.12 (5)	-0.22	0.01	<0.01	0.17 (4)
	Yes, after permission	0.12	0.10	0.24		0.24	0.02	<0.01		0.22	0.01	<0.01	
Data adding	No, not possible	-0.40	0.09	<0.01	0.16 (3)	-0.34	0.02	<0.01	0.17 (3)	-0.12	0.01	<0.01	0.10 (5)
	Yes, possible	0.40	0.09	<0.01		0.34	0.02	<0.01		0.12	0.01	<0.01	
Costs	Per €	-0.03	<0.01	<0.01	0.52 (1)	-0.01	<0.01	<0.01	0.26 (2)	-0.01	<0.01	<0.01	0.22 (2)
Class probability model[2]													
Constant		0.42	0.13	<0.01		-0.19	0.16	0.23		-	-	-	
Having one or more chronic diseases		-0.33	0.15	0.03		-0.54	0.18	<0.01		-	-	-	
Class probability													
Average class probability				0.43				0.20				0.37	
Model fit[3,4]													
Number of observations				17316									
Log Likelihood				-9996									
AIC				1.158									
BIC				1.174									
Pseudo-R^2				0.475									

Abbreviations: *SE* standard error; *RI* relative importance; *GP* general practitioner; *AIC* Akaike Information Criterion; *BIC* Bayesian Information Criterion.
Notes: [1]All attributes (except the costs attribute) were effects coded. The costs attribute was included as a linear term, as transformations (log linear, quadratic and square root) did not improve the model fit. We tested for left-right bias, and found it to be insignificant. [2]Class 3 does not have parameters in the class probability model as the parameters of class 1 and 2 are relative to class 3. [3]Note that the pseudo-R^2 is not the same as the R^2 that is used in a linear regression model. A pseudo-R^2 of 0.4-0.5 is equivalent to a R^2 between 0.8 and 0.9 [41]. [4]The AIC, BIC, and pseudo-R^2 for a model with two classes were 1.252, 1.262 and 0.431 respectively. The AIC, BIC, and pseudo-R^2 for a model with four classes were 1.140, 1.160 and 0.483 respectively. The AIC, BIC, and pseudo-R^2 for a model with five classes were 1.124, 1.148, and 0.491 respectively.

The constant for no PHR was large and positive in class 1 indicating that all else equal respondents of this class prefer not to have a PHR. On the contrary, the constant for no PHR was very large and negative in class 3 indicating that all else equal respondents of this class prefer to have a PHR. Respondents of class 2 were somewhere in between the other two classes. Their constant for no PHR was small and positive, indicating that they might prefer to have a PHR if the characteristics meet their preferences. For both the eager and the reluctant adopters, the facilitating party of PHR data storage was the most important attribute, followed by cost. For the refusers, cost was by far the most decisive attribute, followed by the facilitator of data storage.

Across all classes, independent organisations and care providers were the most preferred facilitators of data storage (largest positive WTP estimates, Table 4). Although the government level was not significant for refusers and reluctant adopters, it was significantly preferred over commercial companies by the reluctant adopters. The reluctant adopters preferred a PHR that is tethered with their general practitioner (GP), with an interconnected PHR being the next best alternative. The eager adopters preferred a PHR tethered with their GP as well, although this was followed by a PHR tethered with their hospital. The WTP estimate for the latter was significantly higher than the WTP for the reluctant adopters. Although the interconnected level was not significant for the eager adopters, it was significantly preferred over a standalone PHR. Both reluctant and eager adopters preferred to give permission for anonymized data use by third parties compared to no use of their personal data. Across all classes, having the possibility to add own data was preferred over not having this possibility. The WTP estimate for data adding, however, was highest for the reluctant adopters.

Irrespective of the PHR characteristics and their ranges considered in the DCE, the predicted PHR uptake for the refusers was always below 9%, while the uptake of the eager adopters was predicted to always be above 91%. Amongst the reluctant adopters uptake was highly sensitive to PHR attribute levels. The expected uptake for the worst imaginable PHR (commercial company, standalone, no use of data, not possible to add own data, 95€) was 4%, while the expected uptake for the best imaginable PHR (care provider, tethered with their GP, use of data after permission, adding own data, zero costs) was 68%. On average, over all three classes, the predicted uptake was 35% for the worst PHR (if constructed as above), while it was 52% for the best PHR (if constructed as above).

Table 4 | Willingness-to-pay estimates in euros per year[1].

		Class 1: the refusers		Class 2: the reluctant adopters		Class 3: the eager adopters	
Short name of attribute	**Attribute level**	**WTP**	**95%CI[2]**	**WTP**	**95%CI[2]**	**WTP**	**95%CI[2]**
Constant	No PHR	126	85;166	48	37;58	-676	-759;-593
Facilitator	Commercial company	-18	-31;-5	-77	-91;-63	-81	-97;-65
	Independent org. or platform	14	3;24	41	33;49	45	35;55
	Government	1	-11;13	-6	-13;1	-17	-26;-8
	Care providers or health facility	3	-7;14	42	35;48	53	43;62
Connectivity	Standalone	3	-7;14	-26	-35;-18	-51	-64;-39
	Tethered with GP	-5	-17;7	23	16;29	33	23;43
	Tethered with hospital	10	-0.3;20	-14	-27;-7	15	5;24
	Interconnected	-8	-20;3	17	11;24	3	-4;10
Use of data	No, never	-4	-11;3	-22	-25;-19	-37	-41;-32
	Yes after permission	4	-3;11	22	19;25	37	32;41
Data adding	No, not possible	-14	-21;-7	-32	-36;-28	-20	-24;-17
	Yes, possible	14	7;21	32	28;36	20	17;24

Abbreviations: *WTP* willingness-to-pay; *CI* confidence interval; GP general practitioner.
Notes: [1]WTP estimates are based on effects coded levels. [2]95% confidence intervals were calculated using the Delta method [51].

DISCUSSION

We identified three classes of potential PHR users with different preference structures: those who prefer not to have a PHR ("refusers", average class probability 43%), those who prefer to have a PHR ("eager adopters", 37%), and those who prefer a PHR only if the tool is designed in accordance with their preferences ("reluctant adopters", 20%). Those with one or more chronic diseases had a higher probability of belonging to the class of eager adopters. The facilitating party for storage of PHR was the most decisive aspect for the eager and the reluctant adopters, while the costs were most decisive for the refusers. Across all classes, care providers or facilities, and independent organisations were the most preferred facilitators of data storage. The predicted uptake for the reluctant adopters ranged from 4% in case of worst PHR to 68% in case of a PHR with the best attribute levels. The predicted uptake for the refusers was always below 9%, while it was always above 91% for the eager adopters.

Our study, like earlier studies [14, 25, 26, 29, 32, 33], showed that privacy concerns are a barrier to the adoption of PHRs. Privacy thus outranks the potential improvement of quality of care that PHRs yield. The facilitators of data storage that respondents preferred most are the healthcare providers or health facilities and independent organisations or platforms. When a commercial organisation or government would store their data, this would reduce the willingness to use a PHR.

The finding that the chronically ill had a higher probability of belonging to the eager adopters of PHRs might be explained by the fact that chronically ill individuals interact frequently with multiple care providers and need to rely on their cooperation to get the best possible care. This can only be realized when all healthcare providers have access to up-to-date information on the care that was delivered to their patient by other providers. Use of PHRs also promotes sharing of information among healthcare providers and as such enable patient-centred care [52, 53]. This could be achieved best by an interconnected PHR. However, we found that the eager adopters prefer a PHR that is tethered to either their GP or their hospital system. This result can possibly be interpreted as respondents being hesitant towards such integration of data because of privacy reasons.

Study limitations

This study has a number of limitations. First, as in all DCEs, the number of attributes and levels that can be included is limited. We aimed to include all relevant attributes and levels by carefully studying the literature, interviewing experts, conducting focus group discussions with potential PHR users and pilot testing the DCE. However, we cannot exclude the option that we missed an attribute or a level that would have affected our results. Second, although we obtained enough power for the analysis in this paper, the response

rate was low. We were unfortunately not able to track how many respondents read the invitation, or started but did not complete the survey. We expect that the reason for the low completion rate is that panel members of NPCF are not used to this type of questions and that no reminders were sent. Those panel members that were not interested in PHRs might not have accessed the survey. It could therefore be argued that the size of class 1 may be even bigger than estimated and therefore, that the majority of the Dutch population at present is not interested in using a PHR. Third, our sample may not be fully representative for the Dutch general adult population. Respondents were older, more highly educated, and more often had chronic diseases. This pattern is similar to earlier surveys that used the same panel [18, 54]. We performed sensitivity analyses to test the robustness of our findings with respect to uptake in subsamples that are representative for the general population based on age, education or health status (Additional file 2). Except for the uptake of the refusers in the representative sample based on health status, predicted uptakes are relatively stable. Despite the fact that our sample is not representative for the general population, it is an interesting sample because of the overrepresentation of people with a chronic illness.

Implications of the study

The first practical implication of this study is that it is of importance that policy makers and PHR producers target their information campaign to chronically ill people given our finding that those respondents had a higher probability to belong to the eager adopters. This large class of respondents (more than one third of the sample) showed to have great interest in PHRs, irrespective of its characteristics, with an uptake that was predicted to always be above 91%. Given that our sample consisted of people that currently do not have a PHR, they might be willing to have a PHR, but they might not know how to get one, or the PHRs that are currently on the market are not interesting to them. The reasons for this might be an area for further research.

The second practical implication is that it appears to be extremely difficult to increase the uptake of PHRs by creating a better product given our finding that the PHR uptake of only 20.0% of respondents was influenced by the characteristics of a PHR. The ideal PHR of this group would be a PHR, which is facilitated by the care provider and tethered with the GP system. The data would be used for other purposes after permission only and adding own data would be possible at zero costs to the user. There were no incongruences across reluctant and eager adopters with respect to the best imaginable PHR. Policymakers who aim to expand the use of PHRs will be most successful when healthcare providers and health facilities, or independent organizations facilitate the storage of PHR data, while refraining to include market parties in storage of the data. Low costs, some form of connection with other systems, and the option to upload own data are valued positively by potential users, but our results suggest that these aspects will only affect uptake marginally. Producers of PHRs need to convince the potential users that they secure the privacy of the PHR data.

CONCLUSIONS

More than one third of potential PHR users indicate great interest in a PHR, irrespective of its characteristics. Policymakers who aim to expand the use of PHRs will be most successful when healthcare providers and health facilities, or independent organizations facilitate the storage of PHR data, while refraining to include market parties in storage of the data.

REFERENCES

1. Connecting for health. A Public-Private Collaborative. Final report. The Personal Health Working Group, Markle Foundation, 2003. Available from: http://www.providersedge.com/ehdocs/ehr_articles/The_Personal_Health_Working_Group_Final_Report.pdf.
2. Tang PC, Ash JS, Bates DW, et al. Personal health records: definitions, benefits, and strategies for overcoming barriers to adoption. J Am Med Inform Assoc. 2006;13:121-6.
3. Pagliari C, Detmer D, Singleton P. Potential of electronic personal health records. BMJ. 2007;335:330-3.
4. Hess R, Bryce CL, Paone S, et al. Exploring challenges and potentials of personal health records in diabetes self-management: implementation and initial assessment. Telemed J E Health. 2007;13:509-17.
5. Shekelle PG, Morton SC, Keeler EB. Costs and benefits of health information technology. Evid Rep Technol Assess. 2006:1-71.
6. Sidorov J. It Ain't Necessarily So: The Electronic Health Record And The Unlikely Prospect Of Reducing Health Care Costs. Health Affairs (Millwood). 2006;2:1079-85.
7. Kaelber D, Pan EC. The Value of Personal Health Record (PHR) Systems. AMIA Annu Symp Proc. 2008:343–7.
8. Chaudhry B, Wang J, Wu S, et al. Systematic review: impact of health information technology on quality, efficiency, and costs of medical care. Ann Intern Med. 2006;144:742-52.
9. Randeree E. Exploring technology impacts of Healthcare 2.0 initiatives. Telemed J E Health. 2009;15:255-60.
10. [Patient information]. Patiënteninformatie - Informatievoorziening rondom de patiënt. Den Haag: Council for Public Health and Health Care (RVZ), 2014. Available from: https://www.raadrvs.nl/uploads/docs/Advies_Patienteninformatie.pdf. *In Dutch.*
11. Pearce C, Bainbridge M. A personally controlled electronic health record for Australia. J Am Med Inform Assoc. 2014;21:707-13.
12. [Overview of Dutch PHR-platforms]. Notitie Overzicht Nederlandse PGD-platforms. Utrecht: Dutch Federation of Patients and Consumer Organizations (NPCF), 2014. Available from: https://www.npcf.nl/Documenten/PGD/Notitie-Nederlands-Overzicht-PGD-Platforms.pdf. *In Dutch.*
13. [Overview of international PHR platforms]. Overzicht buitenlandse PGD platforms. Utrecht: Dutch Federation of Patients and Consumer Organizations (NPCF), 2014. Available from: https://www.npcf.nl/Documenten/PGD/Notitie_Overzicht_van_buitenlandse_PGD-platforms_DEF.pdf. *In Dutch.*
14. Patel VN, Abramson E, Edwards AM, et al. Consumer attitudes toward personal health records in a beacon community. Am J Manag Care. 2011;17:e104-20.
15. Kahn JS, Aulakh V, Bosworth A. What it takes: characteristics of the ideal personal health record. Health Affairs (Millwood). 2009;28:369-76.
16. Tsai J, Rosenheck RA. Use of the internet and an online personal health record system by US veterans: comparison of Veterans Affairs mental health service users and other veterans nationally. J Am Med Inform Assoc. 2012;19:1089-94.
17. Murphy-Abdouch K. Patient Access to Personal Health Information: Regulation vs. Reality. Perspect Health Inform Manag. 2015;12:1c.
18. Van der Steen J, Van Haastert C. [Report 'Personal Health Record' Experiences with and whishes for Personal Health Records]. Rapport 'Persoonlijk GezondheidsDossier' Ervaringen en wensen met betrekking tot het Persoonlijk GezondheidsDossier. Utrecht: Dutch Federation of Patients and Consumer Organizations (NPCF), 2015. Available from: https://www.npcf.nl/Documenten/meldacties/Rapport_PGD_definitief.pdf. *In Dutch.*
19. Johansen MA, Henriksen E. The evolution of personal health records and their role for self-management: a literature review. Stud Health Technol Inform. 2014;205:458-62.
20. Goldzweig CL, Orshansky G, Paige NM, et al. Electronic patient portals: evidence on health outcomes, satisfaction, efficiency, and attitudes: a systematic review. Ann Intern Med. 2013;159:677-87.

21. Ammenwerth E, Schnell-Inderst P, Hoerbst A. The impact of electronic patient portals on patient care: a systematic review of controlled trials. J Med Internet Res. 2012;14:e162.
22. Jilka SR, Callahan R, Sevdalis N, et al. "Nothing About Me Without Me": An Interpretative Review of Patient Accessible Electronic Health Records. J Med Internet Res. 2015;17:e161.
23. Kalra D, Fernando B. A review of the empirical evidence of the healthcare benefits of personal health records. Yearb Medical Inf. 2013;8:93-102.
24. Tenforde M, Jain A, Hickner J. The value of personal health records for chronic disease management: what do we know? Fam Med. 2011;43:351-4.
25. Cocosila M, Archer N. Perceptions of chronically ill and healthy consumers about electronic personal health records: a comparative empirical investigation. BMJ open. 2014;4:e005304.
26. Tieu L, Sarkar U, Schillinger D, et al. Barriers and Facilitators to Online Portal Use Among Patients and Caregivers in a Safety Net Health Care System: A Qualitative Study. J Med Internet Res. 2015;17:e275.
27. Miller H, Vandenbosch B, Ivanov D, et al. Determinants of personal health record use: a large population study at Cleveland Clinic. J Healthc Inf Manag. 2007;21:44-8.
28. Lee G, Park JY, Shin SY, et al. Which Users Should Be the Focus of Mobile Personal Health Records? Analysis of User Characteristics Influencing Usage of a Tethered Mobile Personal Health Record. Telemed J E Health. 2016;22:419-28.
29. Lafky DB, Horan TA. Personal health records: Consumer attitudes toward privacy and security of their personal health information. Health Informatics J. 2011;17:63-71.
30. Irizarry T, DeVito Dabbs A, Curran CR. Patient Portals and Patient Engagement: A State of the Science Review. J Med Internet Res. 2015;17:e148.
31. Archer N, Fevrier-Thomas U, Lokker C, et al. Personal health records: a scoping review. J Am Med Inform Assoc. 2011;18:515-22.
32. Kruse CS, Argueta DA, Lopez L, et al. Patient and provider attitudes toward the use of patient portals for the management of chronic disease: a systematic review. J Med Internet Res. 2015;17:e40.
33. Luque AE, van Keken A, Winters P, et al. Barriers and Facilitators of Online Patient Portals to Personal Health Records Among Persons Living With HIV: Formative Research. JMIR Res Protoc. 2013;2:e8.
34. Fernandez-Aleman JL, Seva-Llor CL, Toval A, et al. Free Web-based personal health records: an analysis of functionality. J Med Sys. 2013;37:9990.
35. Ryan M. Discrete choice experiments in health care. BMJ. 2004;328:360-1.
36. Clark MD, Determann D, Petrou S, et al. Discrete Choice Experiments in Health Economics: A Review of the Literature. Pharmacoeconomics. 2014;32:883-902.
37. Luce D, Tukey J. Simultaneous conjoint measurement: a new type of fundamental measurement. J Math Psychol. 1964;1:1-27.
38. McFadden D. Conditional logit analysis of qualitative choice behavior. In: Zarembka P, editor. Frontiers of Economics. London: Academic Press; 1974. p. 105-42.
39. Manski CF. Structure of Random Utility Models. Theor Decis. 1977;8:229-54.
40. Lancaster KJ. A new approach to consumer theory. J Polit Econ. 1966;74:132-57.
41. Hensher DA, Rose JM, Greene WH. Applied choice analysis: a primer. New York: Cambridge University Press; 2005.
42. Bridges JFP, Hauber AB, Marshall D, et al. Conjoint analysis applications in health--a checklist: a report of the ISPOR Good Research Practices for Conjoint Analysis Task Force. Value Health. 2011;14:403-13.
43. Lancsar E, Louviere J. Conducting discrete choice experiments to inform healthcare decision making: a user's guide. Pharmacoeconomics. 2008;26:661-77.
44. Ryan M, Farrar S. Using conjoint analysis to elicit preferences for health care. BMJ. 2000;320:1530-3.
45. Hiligsmann M, van Durme C, Geusens P, et al. Nominal group technique to select attributes for discrete choice experiments: an example for drug treatment choice in osteoporosis. Patient Pref Adherence. 2013;7:133-9.

46. Reed Johnson F, Lancsar E, Marshall D, et al. Constructing experimental designs for discrete-choice experiments: report of the ISPOR Conjoint Analysis Experimental Design Good Research Practices Task Force. Value Health. 2013;16:3-13.
47. Greene WH, Hensher DA. A latent class model for discrete choice analysis: contrasts with mixed logit. Transport Res B-Meth. 2003;37:681-98.
48. StatLine. Statistics Netherlans (CBS). Available from: http://statline.cbs.nl/Statweb/. [Last accessed 25 03 2016].
49. Fransen MP, Van Schaik TM, Twickler TB, et al. Applicability of internationally available health literacy measures in the Netherlands. J Health Comm. 2011;16 Suppl 3:134-49.
50. Blais A-R, Weber EU. A domain-specific risk-taking (DOSPERT) scale for adult populations. Judgm Decis Mak. 2006;1:33-47.
51. Hole AR. A comparison of approaches to estimating confidence intervals for willingness to pay measures. Health Econ. 2007;16:827-40.
52. Huba N, Zhang Y. Designing patient-centered personal health records (PHRs): health care professionals' perspective on patient-generated data. J Med Sys. 2012;36:3893-905.
53. Reti SR, Feldman HJ, Ross SE, Safran C. Improving personal health records for patient-centered care. J Am Med Inform Assoc. 2010;17:192-5.
54. [Opinion poll 'Primairy Care']. Meldactie 'Huisartsenzorg'. Utrecht: Dutch Federation of Patients and Consumer Organizations (NPCF), 2014. Available from: https://www.npcf.nl/Documenten/meldacties/RapportHuisartsenzorgDEFmrt2014.pdf. *In Dutch.*

ADDITIONAL FILE 1: DCE SURVEY

Example of a choice set

	Personal Health Record A	Personal Health Record B
Facilitating party for storage of PHR data	Commercial company	Government
Level of connectivity	Standalone	Tethered with your hospital
Use of anonymized data by third parties	No, never	Yes, after permission
Option to upload the person's own data into the PHR	No	Yes
The price you pay per year is:	€15	€30

	A	B
What do you prefer?	☐	☐

	A	B	No PHR
What would you choose?	☐	☐	☐

PHR introductory text

There are many different definitions of personal health records (PHRs). It is important that everyone completes this survey with the same definition in mind. We will therefore first describe what is meant with PHRs in this survey.

A PHR is a digital tool that enables you collect, manage and share information about your health, for example, information on diagnosis, treatments, referrals, etc. This information can be obtained from your healthcare providers (e.g., medical specialists, general practitioners, physiotherapists) and / or yourself. Some PHRs include additional functionalities, for example, being able to schedule appointments with your healthcare provider, request medication refills, providing medication reminders, advice and / or information on health and diseases.

In a PHR you control the data. You decide who you provide access to your data and what parts of your PHR this person can see. Persons that were not given access by you cannot see your data. A PHR is complementary to the registration systems used by healthcare providers.

ADDITIONAL FILE 2: SENSITIVITY ANALYSIS OF PREDICTED UPTAKES[1]

		Class 1: the refusers		Class 2: the reluctant adopters		Class 3: the eager adopters		Weighted average over the 3 classes	
		Best (%)	Worst (%)	Best (%)	Worst (%)	Best (%)	Worst (%)	Best (%)	Worst (%)
Total study sample (N=1,433)		4	<1	68	4	99	92	52	35
Representative sample based on age (N=170)	Based on best/worst PHR levels of total study sample	2	<1	69	3	99	83	52	30
	Based on best/worst PHR levels of this sample	3	<1	73	3	99	87	53	31
Representative sample based on education (N=773)	Based on best/worst PHR levels of total study sample	3	<1	68	5	99	92	51	35
	Based on best/worst PHR levels of this sample	3	<1	68	5	99	92	51	35
Representative sample based on health status (N=465)	Based on best/worst PHR levels of total study sample	16	<1	76	2	99	93	56	29
	Based on best/worst PHR levels of this sample	20	<1	78	2	99	93	58	29

Notes: [1]Best and worst PHRs are constructed based on the best and worst PHR levels of the reluctant adopters. The best PHR in the total study sample had the following levels: healthcare provider or health facility, tethered with the system of your general practitioner, use of data after permission, option to upload own data, and zero costs. The worst PHR had the following levels: commercial company, standalone, no use of data, no option to add own data and 95€.

PART III

The future - methodological studies

8

Unforced choice experiments in health economics: implications of using opt-out, neither or status quo alternatives

Domino Determann
Dorte Gyrd-Hansen
G. Ardine de Wit
Esther W. de Bekker-Grob
Ewout W. Steyerberg
Mattijs S. Lambooij
Line Bjørnskov Pedersen

Submitted

ABSTRACT

The literature on which unforced choice format to use in discrete choice experiments (DCEs) is sparse. We investigate the implications of using different unforced choice formats in different settings: 1) opt-out versus neither in a market where there was no status quo (SQ), and 2) including SQ in addition to opt-out in a market with a SQ. A DCE on Dutch citizens' preferences for personal health records served as our case, and 3404 respondents were randomized to the different unforced choice formats. We used mixed logit error component models to estimate preferences. We found that the use of different unforced choice formats affects marginal utilities and welfare estimates and hence the conclusions that will be drawn from the DCE. Only a minority of respondents (14.3%) stated that they interpreted the neither alternative as analysists normally code it (as a zero option). When adding a SQ alternative, this option was chosen in 55.7% of choice sets, while opt-out was chosen in 11.2% of the choice sets. We recommend that future DCE studies use opt-out instead of neither, and that a SQ should be included in addition to opt-out in markets where a SQ exists.

1. INTRODUCTION

Discrete choice experiments (DCEs) are increasingly being used within the health care context to address a wide range of policy questions [1]. It is acknowledged that the hypothetical choice situations in a DCE should represent the real market situation as closely as possible (e.g. [2-5]). If respondents are forced to choose between hypothetical alternatives in a DCE, while they are not in real markets, welfare estimates might be biased and participation rates overestimated [6]. Therefore, DCEs are often presented as unforced choice situations in which respondents are allowed not to choose any of the hypothetical alternatives through inclusion of an opt-out alternative. However, few authors have paid attention to the influence of different unforced choice formats on DCE results.

Three formats are commonly used to ensure unforced choices in DCEs: 1) respondents are given the possibility not to buy the good or not to choose the service (hereafter referred to as opt-out, see for example: [7-9]), 2) respondents are given the possibility to choose 'none of these', or 'neither of the alternatives' (hereafter referred to as neither, see for example: [10-15]), and 3) respondents are given the possibility to choose their current alternative over the hypothetical alternatives (hereafter referred to as status quo (SQ), see for example: [16-20]). The literature on when to use which format is rather sparse [2, 3, 5, 21-25], and does not offer explicit guidance on which format fits a specific situation. A general advice is to specify to respondents what opt-out, neither, or SQ mean. In addition, researchers are encouraged to collect information on respondents' SQ (if applicable) and include this information in the analysis of the choice data.

In this study we aim to investigate the implications of using different unforced choice formats in two different settings: 1) the use of opt-out versus neither in a market where there is no SQ, and 2) including SQ in addition to opt-out in a market where respondents have a SQ. To our best knowledge, there are no published studies that have compared the inclusion of opt-out with neither. In addition, we are aware of only two studies (outside of health economics) that assess the impact of using a SQ alternative versus an opt-out alternative [2, 25]. Both studies found that results differed between the two formats. It was recommended to consider the applicability of format carefully [2], and to choose a format that enhances the realism of the exercise [25]. Hence, using an opt-out in situations where the SQ may be preferred can bias results, just as the opposite may hold true. Our study adds to the findings of [2] and [25] by investigating how the inclusion of a SQ in addition to an opt-out alternative affects DCE results. This has not been studied before although a few studies (one within health economics) have used this approach empirically to elicit preferences [26-29].

The remainder of the manuscript is divided into five sections. Section 2 *'Research questions'* provides a description of the research motivation and research questions. Section 3

'Methods' includes a description of the case study, the design of the study, and the data-collection procedure. In addition, it provides details on the data analysis. Section 4 *'Results'* contains the comparisons between the different unforced choice formats. Section 5 *'Discussion'* concludes with a discussion of the results, the limitations of the study, and recommendations for future studies.

2. RESEARCH QUESTIONS

2.1 Opt-out versus neither

In DCE guidelines, the term opt-out is used to encapsulate both the neither and the opt-out alternative. The only exceptions are the work by [21, 30] where the authors explicitly distinguish between these two formats. Likewise, the opt-out and neither alternatives are usually coded similarly when analysing the data, i.e. as zeros or as missing attribute levels [23]. Although some respondents may interpret choosing neither of the alternatives as a pure opt-out (not to buy the good or use the service), others may interpret neither in terms of a better alternative for which the researcher usually does not know the associated levels. For example, a DCE on patients' hospital choice in the UK [11] included a neither alternative, and respondents were explicitly explained that: *"… choosing 'neither' corresponded to a decision to look either for alternative treatment outside of the NHS, or opting not to have the operation*". This uncertainty surrounding respondents' interpretation leads to challenges when modelling the neither alternative and may bias the DCE results if zero or missing values are used in cases where other levels would have been more appropriate. Previous studies that included a neither alternative also raised this point about multi-interpretability. For example, in a DCE on family planning (FP) and HIV services [31] where a neither alternative was included, it was stated that *"… 'neither' responses could be interpreted either as a choice not to use FP or not to use any service at all. This would obscure the preferences of individuals who would like to use a service, but who find that the alternatives presented are not suitable*". And, in a DCE on preferences for key dimensions of quality of care [32], it was stated that *"… the study may have not clearly specified what a 'no' response meant—whether it indicated seeking care at a private facility or not seeking facility-based care—which may have implications for the interpretation of the constant terms*".

In the first part of this study, we assess empirically whether DCE results are affected by the inclusion of either an opt-out or a neither alternative, and we survey respondents' interpretation of neither. To the extent that choice format does not affect elicited preferences, we conclude that neither is generally interpreted equivalently to opt-out, i.e. as a zero alternative. However, to the extent that the results differ, this is an indicator that the use of neither introduces other interpretations amongst respondents, hereby changing their decision rule and/or choice behaviour. In this case there is no clear alignment between the

respondents' perception of the neither alternative, and the analyst's subsequent analytical approach. In this case, we will argue that opt-out should be preferred over neither in future DCE studies because using a neither alternative entails imputation of attribute values that are likely to be markedly different than those imagined by respondents. Our first research question is:

Research question 1: Does the inclusion of an "opt-out" instead of a "neither" alternative affect DCE results?

2.2 The inclusion of a status quo alternative

In markets where respondents have a SQ, an unforced choice is not necessarily ensured in a choice situation that includes an opt-out alternative [2, 25]. Respondents that prefer their current alternative to the hypothetical alternative(s) and the opt-out alternative do not have the opportunity to make this choice in choice situations where only an opt-out is included. For example, a DCE on patients' preferences for their choice of dentist [7], included two hypothetical alternatives and an opt-out alternative. However, if respondents prefer their current dentist, they are forced to choose one out of three alternatives they would not have chosen in a real market setting, despite that an unforced choice experiment (in its traditional meaning) has been designed. Interestingly, the opt-out alternative is potentially a relevant option for patients who no longer wish to use their current alternative or any of the hypothetical alternatives, as historical decisions do not necessarily reflect future decisions (see e.g. [33] for a study on temporal stability in preferences). Including both an opt-out and a SQ alternative in a DCE has not been explicitly mentioned in the literature as a potential unforced choice format to use, and only a limited number of studies have included both an opt-out alternative and a SQ alternative in each choice set. The only study that we are aware of within health economics is a DCE on preferences for preventive asthma medication in which asthma patients were asked to choose between hypothetical medication, the medication they were currently using, and no medication [26, 27]. The authors argue that the opt-out was used together with the SQ to: "*ensure that respondents were not forced to make a choice between two alternatives, when they might choose neither in practice*". For examples from other fields of economics where a SQ and an opt-out are also included see [28, 29].

In the second part of this study, we assess empirically whether the inclusion of a SQ in addition to an opt-out alternative affects the DCE results in markets where there is a SQ. We argue that if we find differences, this is indicative of the importance of also including a SQ alternative to ensure a truly unforced choice. In this case choice sets that include a SQ in addition to opt-out should be preferred to opt-out only in future DCE studies. Our second research question is:

Research question 2: Does the inclusion of a "status quo" alternative in addition to an "opt-out" alternative affect DCE results in markets where there is a status quo?

3. METHODS

3.1 Case study details

A DCE on Dutch citizens' preferences for personal health records (PHRs) served as our case. A PHR can be defined as: *"An electronic application through which individuals can access, manage and share their health information, and that of others for whom they are authorized, in a private, secure, and confidential environment"* [34]. PHRs are seen as a tool to empower patients [35], give patients more control over their health care process [35], and reduce administrative workload for health care practitioners. Unfortunately, many PHRs that are currently on the market are not based on patients' needs [36], and uptake is limited. Insights in the value that consumers place on various PHR characteristics can help developers to improve their products, and advice policy makers on furthering conditions for the use of PHRs [37].

The case is appropriate for our methodological study since this is a market where some respondents do not have a SQ, while others do. Hence, we allocate respondents who do not have a PHR to DCE questions relating to research question 1, while those with a PHR are presented with DCE scenarios that are designed to answer research question 2.

3.2 Study design

The DCE included five attributes with two to six levels each (Table I). These attributes and levels were carefully selected based on a literature study, interviews with experts (policy makers and researchers; N=5), focus group discussions (N=4, with a total number of 25 participants similar to the target group of the DCE), and brainstorm sessions with co-authors.

By minimizing the D-efficiency criterion, a main-effects efficient Bayesian design based on best guess priors was generated using Ngene software (Choice Metrics, version 1.1.1.). These priors (small positive and negative values) were used to increase the efficiency of the design by avoiding dominant choice sets. We created a design with 12 choice sets to ensure enough degrees of freedom to estimate all main effects (assuming a non-categorical cost attribute), in which also level balance was achieved.

Table I | Attributes and corresponding levels included in the DCE.

Attributes	Attribute levels	Variable name
Facilitating party for storage of PHR data	Commercial company	Reference level
	Independent organization or platform	Stor1
	Government	Stor2
	Health care providers or health facility	Stor3
Level of connectivity	Standalone	Reference level
	Tethered with your GP system	Tech1
	Tethered with your hospital system	Tech2
	Interconnected	Tech3
Use of anonymized data by third parties	No, never	Reference level
	Yes, after permission	Data1
Option to upload the person's own data into the PHR	No, not possible	Reference level
	Yes, possible	Use1
The price you pay per year to make this PHR available to society	€0, €15, €30, €50, €70, €95	Cost

Abbreviations: *GP* general practitioner; *PHR* personal health record.

Choice sets consisted of two unlabelled hypothetical PHR alternatives, and one of the unforced choice formats. We used a split-sample design in which respondents were, given their PHR status, randomized to a certain split based on date of birth (Table II). The opt-out alternative was formulated as 'no PHR', neither was formulated as 'none of these', and the SQ alternative was formulated as 'the PHR that I currently have' (Additional file 1). The same choice sets were used across all splits.

Table II | Overview of splits within the study.

Split	PHR status[1]	Randomization[2]	Choice sets 1-12	Choice sets 12-24	Research question[3]
Split 1	PHR -	50% of PHR -	a, b, opt-out	N/A	1
Split 2	PHR -	50% of PHR -	a, b, neither	N/A	1
Split 3a	PHR +	25% of PHR +	a, b, opt-out	a, b, SQ, opt-out	2
Split 3b	PHR +	25% of PHR +	a, b, SQ, opt-out	a, b, opt-out	2

Abbreviations: *PHR* personal health record; *SQ* status quo; *N/A* not applicable.
Notes: [1]The self-reported PHR status of a respondent; [2]Respondents were, given their PHR status, randomized between splits based on their day of birth (odd or even). The data of 50% of respondents with a PHR was not used for the current study. [3]For research question 1 we used between respondent comparisons, while for analysis of research question 2 we use within respondent comparison. The data of split 3a and split 3b was pooled to accommodate potential ordering effects.

3.3 Survey instrument

The survey started with an introduction to the study and the topic. Next, respondents were screened for their PHR status and given this status subsequently randomized between splits based on their day of birth (odd or even). Those respondents that stated that they currently have a PHR were asked to select the attribute levels that they thought were closest to their

current PHR (the information was used in subsequent analyses). Hereafter, the included attributes and their corresponding levels, as well as the DCE task, were explained to respondents. Subsequently, respondents completed either 12 or 24 choice sets, depending on the split they were randomized to (Table II). The choice sets were followed by an open text question on what respondents had in mind for the neither alternative (split 2 only). All respondents were asked on perceived difficulty of the DCE questions, certainty in their answers to the DCE questions, and their perception of the number of DCE questions. These questions were followed by questions about sociodemographic characteristics. The draft survey was pilot tested using think aloud interviews (N=4) to test for respondent understanding and to improve the wording of the survey. Next, a formal pilot test was conducted (N=51). Pilot test data was analysed using multinomial logit models and estimates were used as priors for the final DCE design.

3.4 Data collection

The survey was administered to Internet panel members of the Dutch Federation of Patient and Consumer Organizations (NPCF) in December 2015 and January 2016. All panel members (N=22,841) received an email with the URL to the online survey and were asked to complete it on a voluntary basis. Data-collection was terminated once the number of respondents completing the survey had decreased to a few respondents per day. Because of practical limitations at the NPCF, no reminders were sent out. Formal testing of the study protocol by a Medical Ethics Committee was not necessary, as a questionnaire study amongst healthy volunteers of an Internet panel does not fall within the scope of the Dutch Act on Medical Research Involving Human Subjects.

3.5 Data analysis

To answer our research questions we performed a number of tests across the relevant splits. In all cases we made comparisons on the number of respondents that always chose opt-out/neither/SQ, respectively, and the total number of times the opt-out/neither/SQ alternative was chosen. Descriptive analyses were performed in SPSS version 22.

To estimate utilities, we used a random utility theory framework where the true but latent utility for alternative *i* of individual *n* can be written as:

$$(1) \qquad U_{in} = V_{in} + \varepsilon_{in}$$

V_{in} represents the observable systematic component of utility, which is the explainable proportion of the choice of alternative *i* of individual *n*, and ε_{in} is the non-explainable proportion representing the unobservable and random treated component. Assuming a linear additive utility function, the observable component for individual *n* for alternative *i*

becomes $V_{in} = \beta X_{in}$ where $X_{in} = (X_1, X_2, ..., X_r)$ is a vector of attributes. The linear predictor, V, of the applied models is shown below.

(2) $$V_{n=} \alpha_n + \beta_1 stor1_n + \beta_2 stor2_n + \beta_3 stor3_n + \beta_4 tech1_n + \beta_5 tech2_n + \beta_6 tech3_n + \beta_7 data1_n + \beta_8 use1_n + \beta_9 cost_n$$

Where α refers to the alternative specific constant (ASC) appropriate for that split. In the split where both a SQ and an opt-out are included, two ASCs were estimated. In that case, α is the sum of the estimates for the two constants.

Assuming that the error terms are independent and identically distributed (iid) extreme value random variables, makes a logit model appropriate. Assuming that the coefficients vary over respondents with density $f(\beta)$, a mixed logit random parameters model can be specified which also takes account of the panel structure in the data and relaxes the independence of irrelevant alternatives (IIA) assumption. The probability of choosing alternative *i* for individual *n* is given by:

(3) $$P_{in} = \int \left(\frac{e^{\mu X_{in} \beta}}{\sum_{j=1}^{J} e^{\mu X_{ij} \beta}} \right) f(\beta) d\beta$$

Where μ is the scale parameter which is inversely related to the error variance [38].

We use an error component specification, by specifying the ASC(s) as random and decompose them into mean β and deviation η_n such that

(4) $$\alpha_n = \beta x_{in} + \eta_n x_{in}$$

All other parameters are held fixed. Error component models were estimated in Stata (version 14), using the 'mixlogit' command using 1000 Halton draws [39].

Individual SQ levels were incorporated in the analysis for research question 2. In all models the attributes (except the cost attribute) were effects coded [7] to enable the interpretation of the estimates for the ASCs. Marginal utilities were calculated afterwards and presented in the tables in the results section. Different functional forms for the cost attribute were tested in all splits (linear, log linear, quadratic and square root transformations), and a linear form was chosen as model fits (AIC and BIC) showed that this was just as good as or better than the other functional forms tested. We only included an ASC for the opt-out/neither/SQ alternative since there was no left-right bias for the hypothetical scenarios in any of the splits.

We compared model fit, signs and significance of utility parameters, relative importance of attributes (RI), and significance of error components between the relevant splits. In addition, we calculated marginal willingness-to-pay estimates (MWTP). T-tests were conducted to test for differences. We also calculated potential market shares to show how the different choice formats affect these. To take account of the random component in the error term, all probabilities were simulated (using 50,000 pseudo-random draws) based on the estimates from the error component models.

3.5.1. Analyses specific to research question 1

In addition to the analyses described above, we compared sociodemographic variables between splits to test whether the randomisation was successful using chi2 tests and independent sample t-tests for equality of means. We also compared respondents' median response times, answers to perceived difficulty of the DCE questions, certainty in their answers to the DCE questions, and their perception of the number of DCE questions. Moreover, we conducted log likelihood ratio tests as proposed by Swait & Louviere [40] to test for equality in utility parameters and scale across samples. Based on the open text question posed to respondents in split 2 regarding what they had in mind when choosing the neither alternative, responses were categorized by one researcher (e.g. as opt-out (no PHR), as a more ideal alternative, etc.), and when there was doubt another researcher verified it.

We estimated two error component models for the neither split. In the first model, neither was coded as zeros for all respondents as is usual coding practice [23, 41]. This model was compared to the opt-out model. In the second model, we added interaction terms with the ASC based on the categorisation of respondents' answers to what they had in mind when choosing neither. This allowed us to investigate the degree of heterogeneity in perceptions of the neither alternative.

4. RESULTS

4.1 Sample of respondents

The DCE was completed by 3404 panel members. Of those, 2986 (88%) had no PHR and were therefore randomized to split 1 or 2 (opt-out and neither split respectively, research question 1), while 418 (12%) had a PHR and were randomized to split 3a-3d. Only the data of respondents randomized to split 3a and 3bd (N=203) were used in the current study. Of these respondents 5% (N=11) had one or more missing current PHR values. Those respondents were excluded from the study. The Additional files 2 and 3 give an overview of number of respondents in the different splits, and their sociodemographic characteristics. Randomization was successful between the opt-out and neither splits on all observed variables (Additional file 3).

4.2 Research question 1 (Opt-out versus neither)

The median response time was slightly higher in the neither split compared to the opt-out split (Table III). The perceived difficulty, certainty in choices and quantity differed between the two splits. More respondents in the neither split found the choice tasks easy to answer, more were certain of their answers, and fewer perceived the number of choices as too high. Fewer respondents in the neither split always chose the neither alternative, while the neither alternative was chosen significantly more often than the opt-out alternative.

4.2.1. Model 1: Opt-out versus neither

All marginal utilities were significantly different from zero for the two splits, and signs were the same across splits (Table III). The facilitating party for storage of PHR data was the most important attribute, while use of anonymized data by third parties was the least important attribute for both splits. For the rest of the attributes, rank orders differed between splits. The ASC for opt-out was relatively smaller than the ASC for neither. However, the standard error, as well as the standard deviation of the ASC, was larger for opt-out than for neither. Pseudo-R^2, AIC and BIC indicated that the opt-out split had a much better model fit than the neither split. The Swait and Louviere log likelihood test showed that marginal utilities in the two models were not equal. MWTPs suggested significant differences between splits for the storage attribute as well as for tech2 (both higher in opt-out split), and the ASCs (higher in neither split). Calculation of potential market shares predicted that the PHR uptake for the best PHR (based on marginal utilities) would be 51% in the opt-out split and 43% in the neither split, whereas for a PHR with the worst attribute levels, uptake would be 35% in the opt-out split and only 14% in the neither split.

Table III | Descriptive statistics and estimation results for research question 1 (comparison between split 1 and split 2, N=2986).

	Split 1: Opt-out (N=1,443)		Split 2: Neither (N=1,543)		
	Median (IQR)		Median (IQR)		P-value
Response time in minutes	21 (16-28)		22 (17-30)		0.003
Survey perception	**N (%)**		**N (%)**		**P-value**
Perceived difficulty					
Very easy	44 (3.1)		81 (5.3)		<0.001
Easy	257 (17.9)		346 (22.6)		
Neutral	600 (41.9)		643 (42.0)		
Hard	449 (31.3)		396 (25.9)		
Very hard	83 (5.8)		64 (4.2)		
Certainty in answers					
Very certain	125 (8.7)		216 (14.1)		<0.001
Certain	503 (35.0)		616 (40.3)		
Neutral	547 (38.1)		499 (32.6)		
Uncertain	224 (15.6)		178 (11.6)		
Very uncertain	37 (2.6)		20 (1.3)		
Quantity of choice sets					
Too low	209 (14.6)		243 (15.9)		0.031
Exactly right	511 (35.6)		598 (39.1)		
Too high	716 (49.9)		689 (45.0)		
Choice behaviour					
Respondents always choosing opt-out / neither	495 (34.3)		474 (30.7)		0.037
Choice sets in which opt-out / neither was chosen	9198 (53.1)		12212 (66.0)		<0.001
Error component models[1]	**Marginal utility (SE)**	**RI (rank)**	**Marginal utility (SE)**	**RI (rank)**	
Stor1	0.880 (0.041)***	0.31 (1)	0.858 (0.045)***	0.39 (1)	
Stor2	0.454 (0.048)***		0.438 (0.050)***		
Stor3	0.905 (0.048)***		0.871 (0.049)***		
Tech1	0.520 (0.046)***	0.18 (3)	0.611 (0.047)***	0.25 (2)	
Tech2	0.381 (0.045)***		0.351 (0.047)***		
Tech3	0.363 (0.045)***		0.390 (0.048)***		
Data1	0.443 (0.025)***	0.15 (4)	0.534 (0.028)***	0.14 (3)	
Use1	0.341 (0.024)***	0.12 (5)	0.410 (0.028)***	0.10 (5)	
Cost	-0.008 (0.000)***	0.25 (2)	-0.010 (0.000)***	0.12 (4)	
ASC opt-out / ASC neither	0.730 (0.210)***	N/A	1.734 (0.097)***	N/A	
	SD (SE)		**SD (SE)**		
ASC opt-out / ASC neither	7.073 (0.296)***	N/A	3.343 (0.104)***	N/A	
Model fit					
Observations	51948		55548		
Respondents	1443		1543		
Log Likelihood (0)	-17078		-15709		
Log Likelihood (Model)	-10021		-11399		
Pseudo-R^2	0.413		0.274		
AIC	20064		22820		
BIC	20161		22918		

Swait and Louviere test			
Equality in utility parameters	$\chi^2_{0.05}(12)$ =21.03; Statistical value = 252.97 Parameter equality rejected, test stops		
Marginal willingness-to-pay	**MWTP (€)**	**MWTP (€)**	**P-value**
Stor1	114	84	0.001
Stor2	59	43	0.047
Stor3	117	86	0.001
Tech1	67	60	0.357
Tech2	49	35	0.038
Tech3	47	38	0.247
Data1	57	53	0.354
Use1	44	40	0.442
ASC opt-out / ASC neither	95	171	0.016
Market shares	**Probability (Opt-out/PHR)**	**Probability (Neither/PHR)**	
Best PHR	49% / 51%	57% / 43%	
Worst PHR	65% / 35%	86% / 14%	

Abbreviations: *IQR* interquartile range; *SE* standard error; *RI* relative importance; *N/A* not applicable; *ASC* alternative specific constant; *SD* standard deviation; *AIC* Akaike Information Criterion; *BIC* Bayesian Information Criterion; *MWTP* marginal willingness-to-pay; *PHR* personal health record.
Notes: [1] *** Statistically significant at the 0.01 level.

4.2.2. Model 2: Interpretation of neither

Respondents interpreted neither in various ways (Table IV, and Additional file 4). Only 14.3% of respondents stated that they had 'no PHR' in mind, while 12.5% of the respondents stated to have another more ideal PHR in mind than the two hypothetical PHRs presented in the choice sets. Some respondents (0.9%) described that they did not think neither represented a real choice. These respondents never chose the neither alternative in any of the choice sets.

The error component model showed significant interactions between the ASC and (most of) the categories of neither interpretation (Table IV). Respondents interpreting the neither as 'no PHR' had significantly more utility associated with choosing the neither alternative compared to those explaining neither as another more ideal alternative. Respondents belonging to the other categories had a significantly lower utility associated with choosing the neither compared to respondents thinking of an ideal PHR.

Table IV | Interpretation of neither alternative (split 2, N=1,543).

Interpretation of the neither alternative	**N (%)**	
Ideal PHR	193 (12.5)	
Opt-out	220 (14.3)	
Neither in their own words	147 (9.5)	
Explanation of decision rule	600 (38.9)	
Neither is not a real choice	14 (0.9)	
Hard to make a choice between PHRs, neither is the easy option to choose	18 (1.2)	
Blank or uninterpretable	351 (22.7)	
Error component model 2[1]	**Utility (SE)**	**SD (SE)**
Stor1	0.858(0.045)***	
Stor2	0.437(0.050)***	
Stor3	0.871(0.049)***	
Tech1	0.611(0.047)***	
Tech2	0.350(0.047)***	
Tech3	0.390(0.048)***	
Data1	0.534(0.028)***	
Use1	0.410(0.028)***	
Cost	-0.010 (0.000)***	
ASC (reference: Ideal PHR)	2.613 (0.162)***	1.678 (0.157)***
ASC No PHR	2.610 (0.598)***	5.248 (0.650)***
ASC Neither in own words	-0.976 (0.258)***	1.318 (0.406)***
ASC Decision rule	-0.735 (0.187)***	0.927 (0.344)***
ASC Not a real choice[2]	-25.618 (4746.873)	0.726 (3932.678)
ASC Hard to make a decision	-2.009 (0.913)**	3.185 (1.209)***
ASC Blank / uninterpretable	-4.114 (0.378)***	5.310 (0.451)***
Model fit		
Observations	55548	
Respondents	1543	
Log Likelihood (0)	-14442	
Log Likelihood (Model)	-11083	
Pseudo-R^2	0.2326	
AIC	22211	
BIC	22416	

Abbreviations: *PHR* personal health record; *SD* standard deviation; *SE* standard error; *ASC* alternative specific constant; *AIC* Akaike Information Criterion; *BIC* Bayesian Information Criterion.
Notes: [1] *** Statistically significant at the 0.01 level, ** Statistically significant at the 0.05 level, * Statistically significant at the 0.1 level. [2]This estimate is highly insignificant since none of these respondents (N=14) ever chose the neither alternative, hereby not contributing to the ASC.

4.3 Research question 2 (The inclusion of a status quo alternative)

The opt-out alternative was chosen less frequently when a SQ alternative was also included in the choice sets, compared to when only opt-out was included (11.2% vs 30.6%, Table V). Moreover, the SQ alternative was chosen in 55.7% of the choice sets in which it was included. A similar pattern was seen for the proportion of respondents that always chose SQ or opt-out.

All marginal utilities were significantly different from zero in the two splits, and signs were the same across splits (Table V). The relative importance and rank orders were comparable, although costs were slightly more important in the opt-out only model. The ASC for opt-out was negative and significant in both splits, which indicates that respondents who are already on the PHR market, all else equal, prefer to have a PHR. The size of the ASC for opt-out was relatively larger in the SQ and opt-out split. In the SQ and opt-out split, the ASC for the SQ was positive, which indicates that respondents, all else equal, prefer their current PHR to a new PHR. Pseudo-R^2, AIC and BIC showed that the model fit was much better for the split in which the SQ was included in addition to the opt-out. MWTPs differed (some at the 10% level) and there were large differences in magnitude on some attributes. With respect to market shares an uptake of 75% (53%) for the best (worst) PHR was predicted in the DCE with opt-out only, while this prediction was 47% (8%) in the DCE with both the opt-out and the SQ.

Table V | Descriptive statistics research question 2 (within respondent comparison, N=192).

	Split 3a and 3b: Opt-out only		Split 3a and 3b: SQ and opt-out		
Choice behaviour	***N (%)***		***N (%)***		***P-value***
Respondents always choosing opt-out	24 (12.5)		11 (5.7)		0.021
Respondents always choosing SQ	N/A		64 (33.3)		<0.001[1]
Choice sets in which opt-out was chosen	705 (30.6)		258 (11.2)		<0.001
Choice sets in which SQ was chosen	N/A		1283 (55.7)		<0.001[1]
Error component model[2]	**Marginal utility (SE)**	**RI (rank)**	**Marginal utility (SE)**	**RI (rank)**	
Stor1	0.881 (0.097)***	0.31 (1)	0.640 (0.128)***	0.38 (1)	
Stor2	0.504 (0.113)***		0.368 (0.145)**		
Stor3	1.166 (0.117)***		0.820 (0.143)***		
Tech1	0.961 (0.111)***	0.25 (2)	0.558 (0.140)***	0.29 (2)	
Tech2	0.891 (0.111)***		0.513 (0.139)***		
Tech3	0.798 (0.109)***		0.377 (0.142)***		
Data1	0.474 (0.057)***	0.13 (4)	0.364 (0.079)***	0.10 (5)	
Use1	0.239 (0.056)***	0.06 (5)	0.384 (0.079)***	0.11 (4)	
Cost	-0.010 (0.001)***	0.25 (3)	-0.009 (0.001)***	0.12 (3)	
ASC opt-out	-2.989 (0.557)***	N/A	-5.602 (1.190)***	N/A	
ASC SQ	N/A	N/A	1.371 (0.442)***	N/A	
	SD (SE)		**SD (SE)**		
ASC opt-out	5.797 (0.641)***	N/A	7.319 (0.860)***	N/A	
ASC SQ	N/A	N/A	5.595 (0.593)***	N/A	
Model fit					
Observations	6912		9216		
Respondents	192		192		
Log Likelihood (0)	-2382		-2649		
Log Likelihood(Model)	-1622		-1423		
Pseudo-R^2	0.319		0.463		
AIC	3267		2872		
BIC	3342		2964		
Marginal willingness-to-pay	**MWTP (€)**		**MWTP(€)**		**P-value**
Stor1	90		71		0.344
Stor2	51		41		0.583
Stor3	119		90		0.233
Tech1	98		62		0.080
Tech2	91		57		0.073
Tech3	81		42		0.043
Data1	48		40		0.496
Use1	24		42		0.128
ASC opt-out	-304		-618		0.079
ASC SQ	N/A		151		<0.001[3]
Market shares	**Probability (Opt-out/PHR)**		**Probability (SQ/Opt-out/PHR)**		
Best PHR	25% / 75%		42% / 11% / 47%		
Worst PHR	47% / 53%		71% / 21% / 8%		

Abbreviations: *SQ* status quo; *N/A* not applicable; *SE* standard error; *RI* relative importance; *ASC* alternative specific constant; *SD* standard deviation; *AIC* Akaike Information Criterion; *BIC* Bayesian Information Criterion; *MWTP* marginal willingness-to-pay; *PHR* personal health record.
Notes: [1] For the difference in proportions between the SQ and opt-out only. [2] *** Statistically significant at the 0.01 level, ** Statistically significant at the 0.05 level. [3] For the difference in WTP between the ASC opt-out and ASC SQ.

5. DISCUSSION

5.1 Research question 1 (opt-out versus neither)

Completing choice sets that include neither seems to be easier for respondents than completing choice sets that include opt-out. This is most likely because respondents have the freedom to interpret the neither alternative in numerous different ways. Only a minority of the respondents' interpret the neither alternative as it is coded by the analyst, namely as an opt-out alternative with zero attribute levels. We found that the marginal utilities and welfare estimates differ between the two splits, and that the model fit was much better for the opt-out split. Based on differences in results between the two splits, we recommend using opt-out instead of neither in future DCE studies, because the neither alternative entails imputation of attribute values (i.e. all zeroes) that are likely to be markedly different from those imagined by the respondents.

5.2 Research question 2 (the inclusion of a status quo alternative)

Both the descriptive data and the error component model show that some respondents may have chosen the opt-out while they actually had their SQ in mind. When respondents are given the opportunity to choose either opt-out or SQ, a large majority of respondents chooses the SQ. In addition, the model fit was better for choice sets including both opt-out and SQ compared to the model with opt-out only. RIs, MWTPs, and market shares differ between the splits; hence, policy recommendations would be very different depending on which format to use. Based on these results, we recommend to include a SQ together with the opt-out in markets with a SQ to ensure a truly unforced choice.

We have analysed the inclusion of SQ in addition to opt-out. However, one may question whether SQ may be a substitute for opt-out. In theory, we would argue for inclusion of both options in order to ensure the ideal unforced choice. Looking at our empirical results, the opt-out was chosen in 11.2% of the choice sets where SQ was also present. We believe that this result is a conservative estimate since PHRs are renewable goods, where consumers are not expected to change preferences on a daily or even monthly basis. If we had used a case of perishable goods, we would expect to observe an even larger rate of opt-out choices, as choices are more fluctuant. In other cases, for example in the case of lifesaving medication where opting out means not surviving, it can be argued that the opt-out alternative can be left out of the choice sets (hereby only including the SQ). To sum up, the relevance of including an opt-out alternative in addition to SQ may depend on the type of good or service (renewable/perishable) and the elasticity of demand (elastic/inelastic).

5.3 Limitations

Our findings should be interpreted with the following limitations in mind. First, because of technical limitations at the NPCF, we were not able to present respondent specific SQ levels in the choice sets and to pivot the attribute levels of the hypothetical alternatives around the SQ (e.g. xx % more expensive than your SQ). Instead, the SQ alternative was framed as "My current PHR" (see Additional file 1). Our finding that SQ should be included in addition to opt-out might have been even more strongly supported if respondents could have seen their individual SQ levels. Second, we did not include a question on how respondents interpreted the opt-out alternative. Although we do not expect this alternative to be interpreted as heterogeneously as the neither alternative, future research could investigate if respondents indeed perceive the opt-out alternative as the option not to buy the good/ choose the service. Third, only 3,404 out of 22,841 panel members completed the survey (15%). We expect that the reason for the low completion rate is that panel members of NPCF are not used to this type of questions and that no reminders were sent. However, as we used a randomized design with a substantial number of respondents in each split, we do not expect this to influence our conclusions.

5.4 Recommendations for future studies

Based on our findings we make two recommendations:

- Use opt-out instead of neither, as neither represents an unspecified alternative that is interpreted differently by respondents, and seldomly as a true opt-out alternative.

- Include both an opt-out and a SQ alternative in markets with a SQ if this best mimics the real market situation. Given that DCE designs are increasingly web-based and interactive, initial SQ questions can easily feed into tailored DCE questions presented to the individual respondent.

We note that our study provides guidance particularly in those DCEs were it is unclear which format to use. We conclude that including either an opt-out option or a SQ option, does not necessarily ensure unforced choice. Hence, health economists engaged in preference elicitation tasks are urged to carefully consider how to design the unforced choice experiment. This study is the first to make a head-to-head comparison of the implications of including neither versus opt-out alternatives as well as the implications of including SQ in addition to opt-out. We show that choice of these design features of DCE matter, and should receive ample attention amongst researchers applying DCEs.

REFERENCES

1. Clark MD, Determann D, Petrou S, et al. Discrete Choice Experiments in Health Economics: A Review of the Literature. Pharmacoeconomics. 2014;32:883-02.
2. Banzhaf MR, Reed Johnson F, Matthews K. Opt-out alternatives and anglers' stated preferences. In: Bennett J, Blamey R (Eds.). The Choice Modelling Approach to Environmental Valuation. Cheltenham, UK: Edward Elgar Publishing Limited; 2001. p 157-77.
3. Batsell RR, Louviere JJ. Experimental analysis of choice. Mark Lett. 1991;2:199-214.
4. Brazell JD, Diener CG, Karniouchina E, et al. The no-choice option and dual response choice designs. Mark Lett. 2006;17:255-68.
5. Ryan M, Skatun D. Modelling non-demanders in choice experiments. Health Econ. 2004;13:397-402.
6. Boxall P, Adamowicz WL, Moon A. Complexity in choice experiments: choice of the status quo alternative and implications for welfare measurement. Aust J Agr Resour Ec. 2009;53:503-19.
7. Bech M, Kjaer T, Lauridsen J. Does the number of choice sets matter? Results from a web survey applying a discrete choice experiment. Health Econ. 2011;20:273-86.
8. Guo N, Marra CA, FitzGerald JM, et al. Patient preference for latent tuberculosis infection preventive treatment: a discrete choice experiment. Value Health. 2011;14:937-43.
9. Mentzakis E, Ryan M, McNamee P. Using discrete choice experiments to value informal care tasks: exploring preference heterogeneity. Health Econ. 2011;20:930-44.
10. Allen JE, Davis AF, Hu W, et al. Residents' willingness-to-pay for attributes of rural health care facilities. J Rural Health. 2015;31:7-18.
11. Burge P, Devlin N, Appleby J, et al. Do patients always prefer quicker treatment? : a discrete choice analysis of patients' stated preferences in the London Patient Choice Project. Appl Health Econ Health Policy. 2004;3:183-94.
12. Hill M, Suri R, Nash EF, et al. Preferences for Prenatal Tests for Cystic Fibrosis: A Discrete Choice Experiment to Compare the Views of Adult Patients, Carriers of Cystic Fibrosis and Health Professionals. J Clin Med. 2014;3:176-90.
13. Kiiskinen U, Suominen-Taipale AL, Cairns J. Think twice before you book? Modelling the choice of public vs private dentist in a choice experiment. Health Econ. 2010;19:670-82.
14. Li J, Scott A, McGrail M, et al. Retaining rural doctors: doctors' preferences for rural medical workforce incentives. Soc Sci Med. 2014;121:56-64.
15. Linley WG, Hughes DA. Decision-makers' preferences for approving new medicines in Wales: a discrete-choice experiment with assessment of external validity. Pharmacoeconomics. 2013;31:345-55.
16. Becker K, Zweifel P. Age and choice in health insurance: evidence from a discrete choice experiment. Patient. 2008;1:27-40.
17. Brau R, Lippi Bruni M. Eliciting the demand for long-term care coverage: a discrete choice modelling analysis. Health Econ. 2008;17:411-33.
18. Grindrod KA, Marra CA, Colley L, et al. Pharmacists' preferences for providing patient-centered services: a discrete choice experiment to guide health policy. Ann Pharmacother. 2010;44:1554-64.
19. Pedersen LB, Kjær T, Kragstrup J, et al. General practitioners' preferences for the organisation of primary care: a discrete choice experiment. Health Policy. 2012;106:246-56.
20. Salkeld G, Ryan M, Short L. The veil of experience: do consumers prefer what they know best? Health Econ. 2000;9:267-70.
21. Lancsar E, Louviere J. Conducting discrete choice experiments to inform healthcare decision making: a user's guide. Pharmacoeconomics. 2008;26:661-77.
22. Carson RT, Louviere JJ, Anderson DA, et al. Experimental analysis of choice. Mark Lett. 1994;5:351-67.
23. Hensher DA, Rose JM, Greene WH. Applied choice analysis: a primer. New York: Cambridge University Press; 2005.

24. Bateman IJ, Carson RT, Day B, et al. Economic valuation with stated preference techniques: a manual. Cheltenham, UK: Edward Elgar Publishing; 2002.
25. Kontoleon A, Yabe M. Assessing the impacts of alternative 'opt-out'formats in choice experiment studies: consumer preferences for genetically modified content and production information in food. J Agric Policy Res. 2003;5:1-43.
26. Lancsar EJ, Hall JP, King M, et al. Using discrete choice experiments to investigate subject preferences for preventive asthma medication. Respirology. 2007;12:127-36.
27. King MT, Hall J, Lancsar E, et al. Patient preferences for managing asthma: results from a discrete choice experiment. Health Econ. 2007;16:703-17.
28. Czajkowski M, Buszko-Briggs M, Hanley N. Valuing changes in forest biodiversity. Ecol Econ. 2009;68:2910-7.
29. Pedersen LB, Kiil A, Kjær T. Soccer attendees' preferences for facilities at the Fionia Park Stadium: An application of the discrete choice experiment. J Sports Econ. 2011;12:179-99.
30. De Bekker-Grob EW, Ryan M, Gerard K. Discrete choice experiments in health economics: a review of the literature. Health Econ. 2012;21:145-72.
31. Michaels-Igbokwe C, Lagarde M, Cairns J, et al. Designing a package of sexual and reproductive health and HIV outreach services to meet the heterogeneous preferences of young people in Malawi: results from a discrete choice experiment. Health Econ Rev. 2015;5:9.
32. Honda A, Ryan M, van Niekerk R, et al. Improving the public health sector in South Africa: eliciting public preferences using a discrete choice experiment. Health Policy Plan. 2015;30:600-11.
33. Doiron D, Yoo HI. Temporal Stability of Stated Preferences: The Case of Junior Nursing Jobs. Health Econ. 2016; [Epub ahead of print].
34. Connecting for health. A Public-Private Collaborative. Final report. The Personal Health Working Group, Markle Foundation, 2003. Available from: http://www.providersedge.com/ehdocs/ehr_articles/The_Personal_Health_Working_Group_Final_Report.pdf.
35. Randeree E. Exploring technology impacts of Healthcare 2.0 initiatives. Telemed J E Health. 2009;1:255-60.
36. Johansen MA, Henriksen E. The evolution of personal health records and their role for self-management: a literature review. Stud Health Technol Inform. 2014;205:458-62.
37. Tang PC, Ash JS, Bates DW, et al. Personal health records: definitions, benefits, and strategies for overcoming barriers to adoption. J Am Med Inform Assoc. 2006;13:121-6.
38. Train KE. Discrete choice methods with simulation. New York: Cambridge University Press; 2009.
39. Hole AR. Estimating mixed logit models using maximum simulated likelihood. Stata Journal. 2007;7:388-401.
40. Swait J, Louviere J. The role of the scale parameter in the estimation and comparison of multinomial logit models. J Marketing Res. 1993:305-14.
41. Chen LC, Cheng LJ, Zhang Y, et al. Acupuncture or low frequency infrared treatment for low back pain in Chinese patients: a discrete choice experiment. PLoS One. 2015;10:e0126912.

ADDITIONAL FILE 1: CHOICE SET EXAMPLES

We used a dual response design in which respondents were forced to make a choice between two personal health record (PHR) alternatives first, and they were then asked to make an unforced choice. We believed that this approach was necessary given the low expected uptake of PHR.

Opt-out:

	Personal Health Record A	Personal Health Record B
Who facilitates data storage:	Commercial company	Government
Technical aspect of data storage:	Standalone	Tethered with your hospital
Use of anonymized data to improve quality of Dutch health care system	No, never	Yes, after permission
Adding of own data possible:	No	Yes
PHR costs for society. The price you pay per year is:	€15	€30

	A	B
What do you prefer?	☐	☐

	A	B	No PHR
What would you choose?	☐	☐	☐

Neither:

	Personal Health Record A	Personal Health Record B
Who facilitates data storage:	Commercial company	Government
Technical aspect of data storage:	Standalone	Tethered with your hospital
Use of anonymized data to improve quality of Dutch health care system	No, never	Yes, after permission
Adding of own data possible:	No	Yes
PHR costs for society. The price you pay per year is:	€15	€30

	A	B
What do you prefer?	☐	☐

	A	B	None of these
What would you choose?	☐	☐	☐

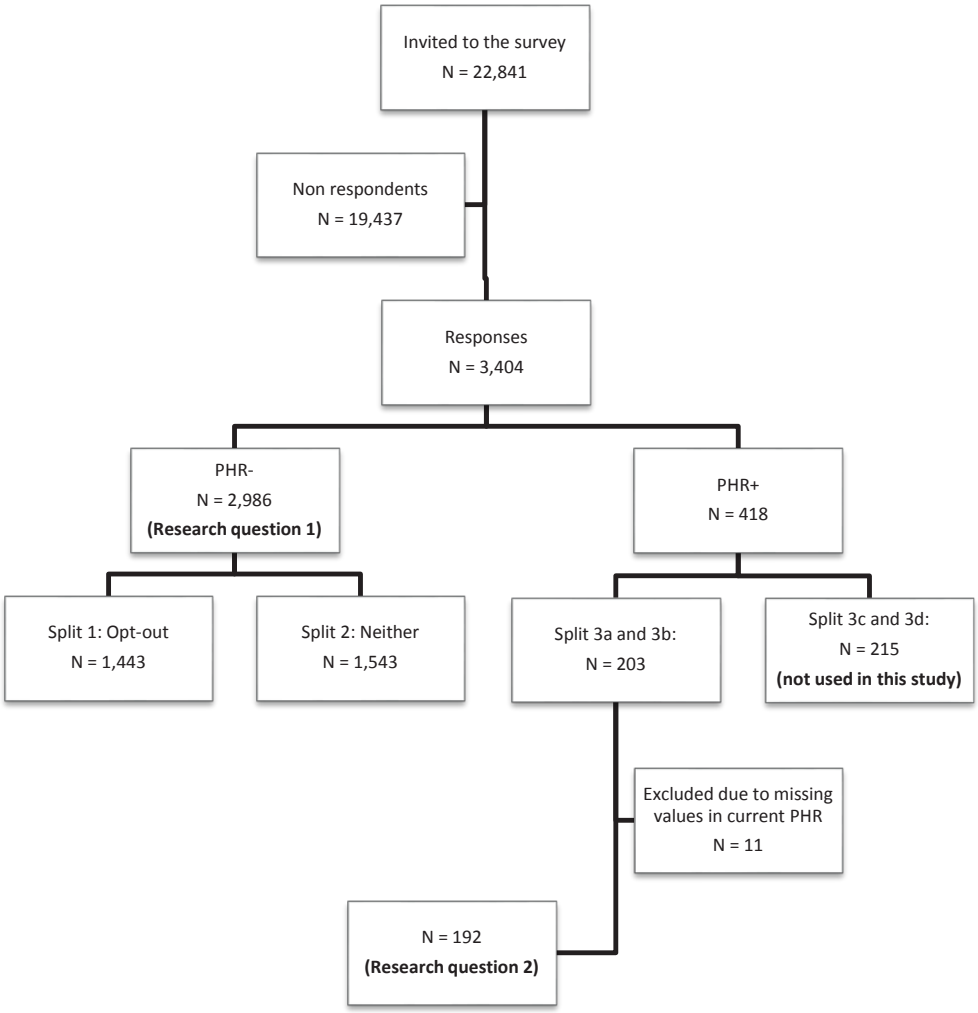
Status quo and opt-out:

	Personal Health Record A	Personal Health Record B
Who facilitates data storage:	Commercial company	Government
Technical aspect of data storage:	Standalone	Tethered with your hospital
Use of anonymized data to improve quality of Dutch health care system	No, never	Yes, after permission
Adding of own data possible:	No	Yes
PHR costs for society. The price you pay per year is:	€15	€30

	A	B
What do you prefer?	☐	☐

	A	B	Current PHR	No PHR
What would you choose?	☐	☐	☐	☐

ADDITIONAL FILE 2: FLOWCHART



Abbreviation: *PHR* personal health record.

ADDITIONAL FILE 3: DESCRIPTIVE STATISTICS

Research question 1		**Split 1: Opt-out (N=1,443)** Mean (SD)	**Split 2: Neither (N=1,543)** Mean (SD)	P-value
Age		61.41 (11.26)	61.75 (11.00)	0.425
		N (%)	**N (%)**	
Gender	Female	698 (51.0)	736 (50.2)	0.677
	Male	671 (49.0)	730 (49.8)	
Education[1]	Low	254 (18.1)	309 (20.6)	0.213
	Average	443 (31.6)	466 (31.1)	
	High	706 (50.3)	722 (48.2)	
Work status	Paid work	992 (70.0)	1090 (71.8)	0.297
	No paid work	425 (30.0)	429 (28.2)	
Income[2]	Low	370 (33.7)	436 (36.3)	0.360
	Average	326 (29.7)	355 (29.5)	
	High	402 (36.6)	411 (34.2)	
Health status	Healthy	325 (22.5)	384 (24.9)	0.129
	Chronically ill	1118 (77.5)	1159 (75.1)	
Health literacy[3]	Low	29 (2.0)	36 (2.4)	0.535
	Adequate	1395 (98.0)	1481 (97.6)	

Research question 2		Split 3a and 3b (N=192)
		Mean (SD)
Age	In years	63.22 (10.65)
		N (%)
Gender	Female	82 (45.6)
	Male	98 (54.4)
Education[1]	Low	54 (29.3)
	Average	51 (27.7)
	High	79 (42.9)
Work status	Paid work	35 (18.5)
	No paid work	154 (81.5)
Income[2]	Low	62 (40.5)
	Average	45 (29.4)
	High	46 (30.1)
Health status	Healthy	26 (13.5)
	Chronically ill	166 (86.5)
Health literacy[3]	Low	11 (5.7)
	Adequate	181 (94.3)
Difficulty	Very easy	5 (2.6)
	Easy	40 (21.1)
	Neutral	68 (35.8)
	Hard	60 (31.6)
	Very hard	17 (8.9)
Certainty	Very certain	14 (7.3)
	Certain	77 (40.3)
	Neutral	62 (32.5)
	Uncertain	33 (17.3)
	Very uncertain	5 (2.6)
Quantity	Too low	24 (12.8)
	Exactly right	63 (33.5)
	Too high	101 (53.7)
		Median (IQR)
Response time	In minutes	28 (21-37)

Abbreviations: *SD* standard deviation; *IQR* interquartile range.
Notes: [1]Educational level was categorized into three groups: low (primary education and lower secondary education), average (higher secondary education and intermediate vocational education), and high (tertiary education). [2]Monthly nett income was categorized into three groups: low (less than €2000), average (€2000-€3000), high (€3000 or more). [3]Subjective health literacy was measured based on the validated Dutch questions of the Set of Brief Screening Questions (SBSQ-D) of Chew [1].

REFERENCE

1. Fransen MP, Van Schaik TM, Twickler TB, et al. Applicability of internationally available health literacy measures in the Netherlands. J Health Comm. 2011;16Suppl3:134-49.

ADDITIONAL FILE 4: QUOTES OF INTERPRETATION OF THE NEITHER ALTERNATIVE

"The elements that I preferred were mixed between PHR A and PHR B, so none of the alternatives offered exactly what I would want. So I thought of a combination where I myself could decide on the options." *[Respondent 1363: ideal PHR]*

"That there was a better alternative" *[Respondent 1815: ideal PHR]*

"The combination of elements did not meet my preferences. I do want a PHR, but only one that has all my elements of choice." *[Respondent 3157: ideal PHR]*

"No PHR" *[Respondent 443: opt-out]*

"Neither is not really a choice" *[Respondent 1610: not a real choice]*

"I like some elements of both PHR alternatives, so I could not make a clear choice" *[Respondent 1003: hard to make a choice]*

9

Impact of survey administration mode on the results of a health related discrete choice experiment: online and paper comparison

Domino Determann
Mattijs S. Lambooij
Ewout W. Steyerberg
Esther W. de Bekker-Grob
G. Ardine de Wit

Submitted

ABSTRACT

Objectives: Electronic data-collection is increasingly used for discrete choice experiments (DCEs). We aimed to study whether paper or electronic administration results in measurement effects.

Methods: Respondents were drawn from the same sample frame (an Internet panel) and completed a nearly identical DCE survey either online or on paper during the same period. A DCE on preferences for basic health insurance served as a case study. We used panel mixed logit models for the analysis.

Results: In total, 898 respondents completed the survey: 533 respondents completed the survey online, while 365 respondents returned the paper survey. There were no significant differences with respect to sociodemographic characteristics between the respondents in both samples. The median response time was shorter for the online sample compared to the paper sample, and a smaller proportion of respondents from the online sample was satisfied with the number of choice sets. In addition, a larger proportion of respondents from the online sample, compared to the paper sample, failed the monotonicity axiom. We found some significant between-sample differences in choice probabilities for the last five choice sets. Although some WTP estimates were higher for the online sample, the elicited preferences for basic health insurance characteristics were similar between both modes of administration.

Conclusions: We find no indication that online surveys yield inferior results compared to paper based surveys, while the price per respondent is lower for online surveys. However, researchers might want to include fewer choice sets per respondent when collecting DCE data online. More research is needed to support our findings.

INTRODUCTION

Discrete choice experiments (DCEs) can be administered in various manners, examples include paper surveys (either completed on site or distributed via mail) and electronic surveys (either on site using a laptop, or distributed via Internet) [1]. Although paper surveys were the most common DCE data-collection mode up to 2009, there has since then been a shift towards electronic administration [2]. Advantages of electronic data-collection over more traditional administration modes include a rapid collection of data, fewer errors in the process of data entry and, potentially, lower costs. Technical possibilities inherent to electronic nature of surveys offer additional benefits, such as being able to track the response time, to randomize or skip questions based on respondents' previous answers and to force respondents to answer all questions. The main concerns regarding electronic data-collection are the representativeness of samples, and reliability and validity of the obtained data. For a more extensive overview of advantages and disadvantages of electronic data-collection, see for example Ekman & Litton [3], or Gelder et al. [4].

Choice of administration mode may affect the responses to a survey, even when the questions asked are identical [5, 6]. Administration mode effects can be divided into two broad effect categories [7], namely 1) sample composition, or representation effects, and 2) measurement effects, which are the focus of this paper. Sample composition effects relate to who responds to the survey. Measurement effects on the other hand relate to the responses elicited, and include social desirability bias (e.g. in the case when an interviewer is present), and satisficing or a respondent's failure to put in sufficient effort to answer the question. Measurement effects occur when the same respondent would provide different responses to identical worded questions between different survey modes [8]. Research within the field of DCEs has shown that changes in design of the survey, for example attribute framing [9], attribute ordering [10], and the presentation of attributes as either words or graphics [11], have unintended influences on obtained estimates, which might result in biased conclusions. In order to be able to assess the appropriateness of increased electronic data-collection, it is therefore essential to study whether mode of administration (paper versus online administration) also affects DCE results.

In 2011, Lindhjem & Navrud reviewed studies that compared survey modes within environmental stated preference research including several studies that compared paper with online administration of DCEs [7]. The authors summarized that evidence of measurement effects was lacking, mainly because of the lack of experimental control and confounding of measurement and sample composition effects within the reviewed studies [7]. We are aware of one additional published study comparing paper with online DCE administration since this review. This study by Boyle et al., showed that Internet-based welfare estimates were 76% of the mail-based welfare estimates on average [12]. The health-related study

described in our paper is a novel contribution to the already existing literature since we study mode effects using a well-controlled design in which a large number of respondents drawn from the same sample frame (an Internet panel) completed a nearly identical survey either online or on paper during the same period. Additionally, almost all studies on this topic have been conducted in environmental economics. We found only two health-related stated preference studies comparing different administration modes, however not paper versus online. Pieterse et al. [13] used an adaptive conjoint analysis survey on rectal cancer treatment outcomes to study whether valuations differed between administration via a portable computer and administration through Internet; no mode effects were found. Mulhern et al. [14] studied whether mode (online versus computer-assisted personal interview) had an effect on health state valuations; no effect was found.

We aimed to study whether the choice of DCE administration mode (paper versus online) results in measurement effects.

METHODS

Case study details

Our study on administration mode was embedded in a DCE on choices of Dutch consumers with regard to basic health insurance. In the Netherlands, consumers have the chance to switch health insurance on a yearly basis. Within our healthcare system with managed competition, health insurers are supposed to act as prudent buyers of care on behalf of their customers. To fulfil this role adequately, understanding consumer preferences for health plan characteristics is of vital importance. For that reason, we conducted a DCE to quantify consumer trade-offs for basic health plan characteristics. We used this DCE application as a case for the current methodological study.

The DCE included four attributes: (1) level of choice of care provider, (2) the prime focus of insurers' contracting practices, (3) service benefits provided by the insurer, and (4) monthly premium (Table 1). This selection was based on a literature study, interviews with experts in the field of health insurance (researchers, policy makers, and employees of a health insurance company, n=9) and focus group discussions (n=4) with a total number of 23 participants. A subset of all possible choice sets was generated using best-guess fixed priors in Ngene design software (version 1.1.1.) by minimizing the D-efficiency criterion. These priors were based on the literature study and the expert interviews. In addition, experts suggested that it was unrealistic to include combinations of the lowest level of freedom of provider choice with the highest premium level and vice versa. These restrictions were therefore imposed upon the design. To assure level balance and to have enough power to estimate all the parameters, the final design consisted of 30 choice sets, blocked into two

versions. Each choice set consisted of two basic health plan alternatives: alternative A and alternative B. We choose not to include an opt-out alternative, as having a basic health insurance plan is obligatory in the Netherlands.

Table 1 | Attributes and attribute levels included in the DCE.

Attributes	Levels	Explanation to respondents
Choice of care providers	50% (ref) 80% 100%	For fully reimbursed hospital care consumers can go to [level] of all hospitals. If consumers go to another hospital none (in the case of level 50%), 75% (in the case of level 80%) or all (in the case of level 100%) of the health care costs will be reimbursed respectively
Primary focus of provider contracting	Price of care (ref) Quality of care Social responsibility	Health insurers contract care providers mainly focused on either the quality of the care provided, price of the care provided or based on social responsibility
Level of service benefits	Negative ratings on consumer websites (ref) Positive ratings on consumer websites	Customer services include the friendliness and ease of contact with the health insurer and the accurate and prompt reimbursement of the claims of consumers
Premium	€70 €85 €100 €115	The monthly nominal premium for a basic health insurance that needs to be paid directly by the consumer to the health insurer

Survey instrument

The survey contained two parts. Part one included an introduction, an example choice set, and either one of the blocks of 15 choice sets. In addition, we included two extra choice sets per respondent to test two important axioms underlying DCE methodology: (1) the number of respondents that chose the clearly dominant alternative (monotonicity axiom), and (2) the number of respondents that chose the same alternative in repeated choice sets (completeness axiom) were assessed [15-17]. It could be hypothesized that there might be differences in the number of respondents violating these axioms between administration modes. Thus, each respondent needed to complete 17 choice sets in total. Part two contained some additional questions (e.g. on respondents' characteristics and on the survey perception). The draft paper survey was pilot tested using think aloud interviews (n=4) to test for respondents' understanding. As a result, the wording of the survey was improved where necessary. Subsequently, a formal online pilot test (n=56) was conducted. The pilot test data was analysed using multinomial logit models and estimates were used as (fixed) priors of the final DCE design.

The survey instrument was kept as identical between the two modes as practically possible. However, the surveys differ on two points. First, the online survey forced respondents to answer before they could move on to the next question (standard practice in online surveys);

while in the paper survey there obviously was a possibility for item non-response. This also implies that the online survey is not scrollable, while respondents that completed the paper survey could flip through the whole survey. Second, we aimed to report response times for the two samples, because previous research shows a relation between response time and satisficing [18]. Respondents that completed the paper survey were therefore asked to write down the time when they started with the survey as well as to note the time when they finished the survey. This was done automatically for respondents who completed the survey online.

Data-collection

We used the rule of thumb as suggested by Orme [19] to determine the minimum required sample size needed to answer our research question. We aimed to have at least n=900 respondents completing the survey (n=450 per mode of administration). Afterwards, we used the parametric sample size calculation method as suggested by de Bekker-Grob et al. [20] to check whether the statistical power was sufficient, which it was.

A Dutch research company (CG selections) was selected to collect the data for this study among their panel members. Ideally, panel members should have been randomly assigned to either one of the two survey modes. This was not possible however because panel members of this company are used to complete surveys online and not to complete paper surveys. Hence, it would have been unusual for panel members to receive a paper survey at home without a formal invitation to participate in the study. We therefore opted for the second best option, namely to first send a standard invitation email to panel members. When respondents clicked on the URL that was included in the email, they were first asked three questions on their age, gender and educational level, to ensure a representative distribution of the general adult population of The Netherlands in the sample. This was especially important since the aim of our study was to assess measurement effects between modes and not sample composition effects. Panel members who met the quotas for age, gender and educational level, were asked whether they would be willing to fill out a paper survey. Respondents that were willing to complete a paper survey were asked for their postal address and the research company sent the survey to them. Respondents were paid a small incentive once they had returned the completed survey to the research company. Another research company (Datadesk) entered answers given to the paper survey into SPSS. Those that were not willing to complete the paper survey were subsequently asked whether they would be willing to complete the survey online. Respondents that completed the survey online received a small incentive as well. When the target number of panel members that agreed to participate in the paper-based survey was reached, panel members were no longer asked whether they were willing to complete the paper survey, they instead were immediately asked to participate in the online study. The data was collected between December 17th, 2014 and January 14th, 2015.

Ethical statement

Formal testing of the study protocol by a Medical Ethics Committee was not necessary, as a survey study amongst healthy volunteers of an Internet panel does not fall within the scope of the Dutch Act on Medical Research Involving Human Subjects.

Analysis

First, we evaluated first whether respondents of both survey modes were comparable with respect to sociodemographic characteristics using independent t-tests and chi-squared tests in SPSS (version 22).

Second, descriptive results were compared between administration modes. In addition, the number of respondents that fail the monotonicity axiom and completeness axiom were assessed. The number of missing responses in choice sets and the choice probabilities were also analysed.

Third, Nlogit (version 5.0) was used to construct both panel latent class models (LC) and panel mixed logit models (MIXL). These models account for the panel nature of the data (i.e. each respondent completing a number of choice sets), as well as for preference heterogeneity between respondents [21]. A LC model with two classes, in which it was tested whether mode of administration was a significant class predictor, had a lower fit (based on AIC, pseudo-R2, and log likelihood) compared to the final MIXL model and is therefore not reported on in the rest of the paper. In the MIXL models, all attributes (except for the premium attribute) were effects coded. In order to be able to calculate the willingness-to-pay (WTP), the premium attribute was included as a fixed parameter, while it was assumed that all other attribute level estimates were random and follow a normal distribution. An alternative specific constant (ASC) was included in the model to account for the possibility of left-to-right bias. A positive significant constant indicates that respondents were more likely to select the first alternative they read when answering the choice sets (first column, i.e. 'health insurance A'). The final analysis included two models. In the MIXL model on pooled data, it was tested whether the mode of administration interacted with all attribute levels. Because of potential between-sample differences in scale of the (unobserved) variance of the error term, we estimated the same MIXL (except for the interactions) in the two samples separately. The coefficients are still not directly comparable between these two models; however, coefficient significance, as well as signs, relative importance and WTP estimates can be compared. Mean marginal monthly WTP estimates were calculated by taking the ratio of the change in the attribute level of interest over the negative coefficient of the premium attribute. By doing this, the scale factor becomes redundant in the equation. The 95% confidence levels around the mean WTP estimates were approximated using the Delta method [22].

RESULTS

The URL in the email was clicked on by 1313 panel members. Of those 1313 members, 329 (25%) did not meet the quotas for age, gender and/or educational level, and these members were therefore directly excluded from participation. A large majority of the target number of 450 panel members that agreed to participate in the paper survey (n=365, 81%) returned the completed paper survey to the research company. Of the 534 panel members that were asked to participate in the online survey, 533 participated. Thus, a total of 898 respondents completed the survey. As sample completion was steered by quotas based on socio-demographic characteristics, there were no significant differences in sociodemographic characteristics between the respondents of both samples (Table 2).

Table 2 | Demographic characteristics of study population, stratified by administration mode.

		Total sample (n=898)	**Online sample (n=533)**	**Paper sample[1] (n=365)**	**P-value[2]**
Age in years	Mean ± SD (range)	47.7 ± 15.0 (19-74)	47.0 ± 14.9 (19-70)	48.7 ± 14.9 (19-74)	0.12
Age groups	18-29 years (%)	17.8	18.8	16.5	0.40
	30-44 years (%)	24.1	25.3	22.3	
	45-59 years (%)	27.8	27.6	28.2	
	60+ years (%)	30.2	28.3	33.0	
Gender	Female (%)	49.7	50.7	48.2	0.47
Educational level[3]	Low (%)	23.8	24.0	23.4	0.76
	Average (%)	41.3	42.0	40.2	
	High (%)	34.9	34.0	36.4	
Gross income per month[4]	Low (%)	33.3	34.1	32.0	0.72
	Average (%)	20.4	19.9	21.3	
	High (%)	27.9	27.4	28.7	
	Does not know, or does not want to tell (%)	18.3	18.6	18.0	
Working status	Paid, full or part-time (%)	53.1	54.4	51.2	0.35
Number of chronic diseases[5]	Mean ± SD (range)	1.23 ± 1.49 (0-9)	1.21 ± 1.54 (0-9)	1.27 ± 1.41 (0-6)	0.52
	None (%)	41.7	44.7	37.3	0.06
	1 disease (%)	25.4	24.2	27.1	
	2 diseases (%)	16.5	14.1	20.2	
	3 diseases (%)	8.3	8.3	8.3	
	> 3 diseases (%)	8.1	8.8	7.2	
Subjective health literacy[6]	Low health literacy (%)	3.7	4.5	2.5	0.11

Abbreviation: *SD* standard deviation.

Notes: [1]Contains missing data: age and age groups (n=7), gender (n=2), educational level (n=2), income (n=3), number of chronic diseases (n=3); [2]P-value for online vs. paper sample; [3]Educational level was subdivided into three groups: low (primary education and lower secondary education, average (higher secondary education and intermediate vocational education), and high (tertiary education); [4]Gross income was measured on a categorical scale. Respondents' answers were subdivided into four groups: low (less than €2000), average (between €2000 and €3000), high (€3000 or more), and does not know, or does not want to tell; [5]Respondents were asked to select from a predefined list which chronic disease(s) they suffered. 6Subjective health literacy was measured using the Set of Brief Screening Questions (SBSQ-D) of Chew [23].

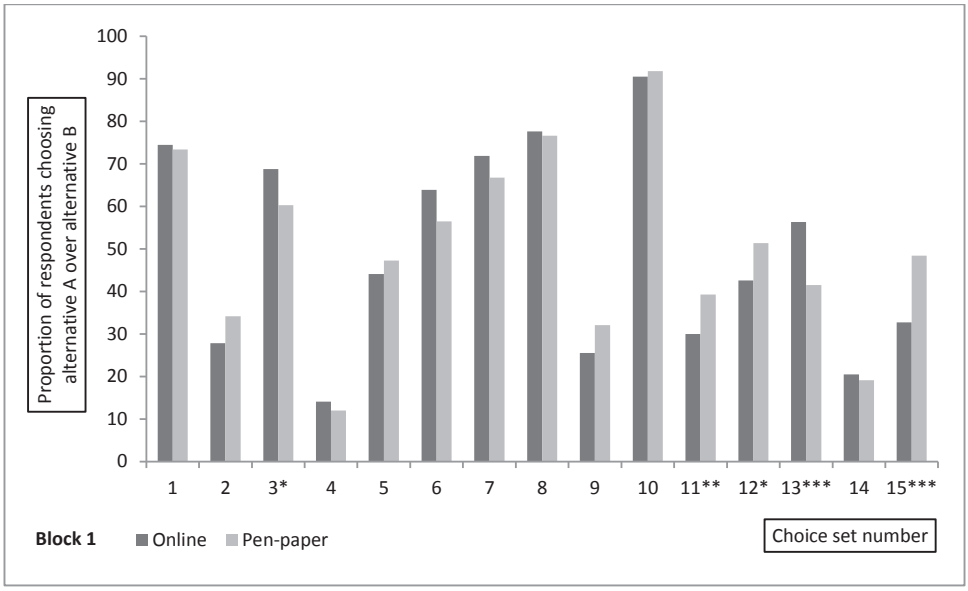
Descriptive statistics

The logged response time was significantly shorter in the online sample compared to the reported response time of the paper sample (median, 14.0 vs. 24.5 minutes, Table 3). The majority of respondents of both samples (65 to 69%) found the survey easy or very easy to complete. The number of choice sets was perceived differently between the samples. Significantly more respondents from the online sample found the number of choice sets too high, compared to respondents from the paper sample (24.0% vs. 13.0%). Almost half of the paper sample (46.6%) was indifferent when asked (after the survey) whether they preferred to complete DCE surveys either online or on paper, and more than one-third (38.6%) stated to prefer paper over online. The majority of the online sample stated that they prefer to complete DCE surveys online (68.5%). A small number of respondents of both samples searched for information during the survey. Of those who did, respondents from the paper sample stated to having discussed the survey with others, while those who completed the survey online stated to having searched for information online. Slightly more respondents of the paper sample, compared to the online sample, chose the dominant alternative in the choice set with a clearly dominant alternative (99.2% vs 97.0%). There was no difference between samples with respect to the completeness axiom. Because answering the choice sets was mandatory in the online sample, there was no missing choice data, while 21 choice observations (0.4%, from n=7 respondents) were missing from the paper sample. The choice probabilities for individual choice sets differ between administration modes, especially in choice sets 10-15 (Figure 1).

Table 3 | Descriptive statistics stratified by administration mode.

		Online sample (n=533)	Paper sample[1] (n=365)	P-value
Response time in minutes	Median (IQ range)[2]	14.0 (10.0-19.0)	24.5 (17.3-31.0)	<0.01
Level of difficulty	Easy or very easy (%)	69.0	65.2	0.20
	Indifferent (%)	23.6	28.8	
	Hard or very hard (%)	7.3	6.0	
Number of choice sets	Too small (%)	6.6	7.2	<0.01
	Just right (%)	69.4	79.8	
	Too high (%)	24.0	13.0	
Mode preference	Paper (%)	3.4	38.6	<0.01
	Online (%)	68.5	14.8	
	Indifferent (%)	28.1	46.6	
Searched for information during survey	Yes (%)[3]	8.8	9.6	0.69
	Discussed with others	25.5	71.4	
	On the Internet	57.4	5.7	
	Own health plan	27.7	22.9	
Monotonicity axiom	Prefers better over worse levels in choice set with a clearly dominant alternative (%)	97.0	99.2	0.03
Completeness axiom	Chose the same alternative when the choice set was repeated later in the survey (%)	82.4	81.6	0.78
Missing choice data	No. of missing responses in choice sets (no. of respondents)	N.A.	21 (7)	N.A.

Abbreviations: *IQ* range Interquartile range; *N.A.* Not applicable.
Notes: [1]Contains missing data: response time (n=5), number of choice sets (n=4), searched for information during survey (n=4); [2]Median and IQ range are reported instead of mean and standard deviation because of a number of outliers (respondents who left the survey open in their browser for a longer period) in the online sample. Note that the response time was logged automatically in the online sample, while respondents from the online sample were asked to write down the begin- and finish time; [3]The sum of the proportions in the online sample is more than 100, because some respondents chose more than option.



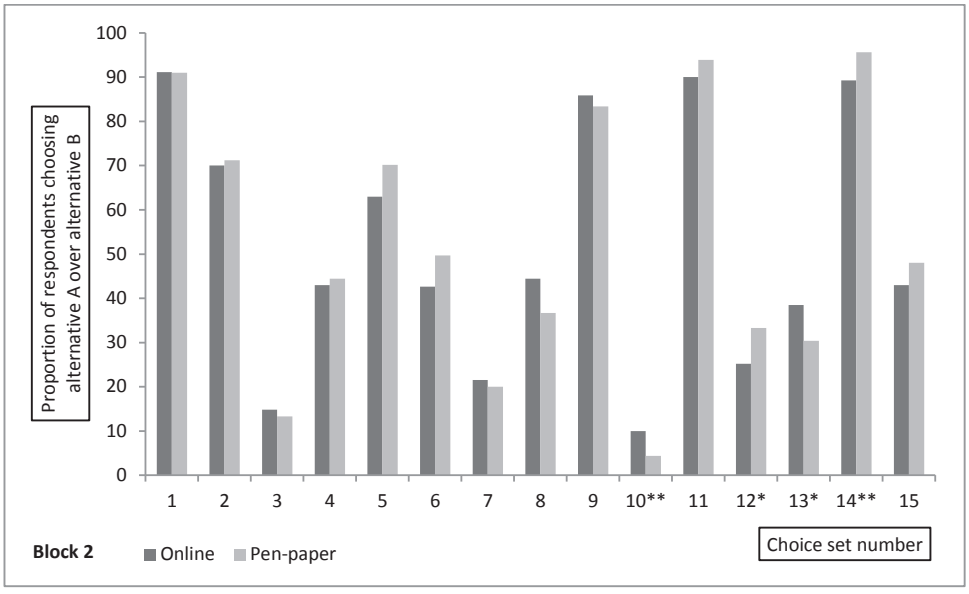


Figure 1 | Choice probabilities.

Mixed logit model results

Mode of administration significantly interacted with two out of four attributes, i.e. the attribute 'primary focus of provider contracting' and the 'monthly premium' (Table 4, column total sample), indicating that the mean estimates differ significantly between the two samples. The interactions between mode and the other attributes, as well as the interaction between mode and the ASC, were not significant.

The separate models of the two samples show that all attribute level mean estimates were statistically significant and had identical signs (Table 4, online sample & paper sample). However, the mean estimate of the ASC was statistically significant in the paper sample only (indicating left-to-right bias), while standard deviations of the ASC were not significant in both samples. The statistically significant attribute level estimates of the standard deviations are an indication of preference heterogeneity, except for the attribute level 'social responsibility'. The model fit was slightly better for the online sample.

The rank order of importance of attributes and importance scores were similar across both samples. Choice of care provider was considered most important, followed by premium per month, level of service benefits, and the primary focus of provider contracting. The importance scores differed slightly between the samples: choice of care providers was slightly more important in the online sample, compared to the paper sample, while premium was more important in the paper sample.

WTP estimates of the choice of care provider attribute were significantly higher in the online sample compared to the paper sample (ratio 1.32) and the level of service benefits (ratio 1.16, Table 5). No significant differences in WTP estimates were found for the attribute levels of primary focus of provider contracting.

Table 4 | Panel mixed logit results[1,2,3,4].

		Total sample (n=898)		Online sample (n=533)			Paper sample (n=365)		
Attributes	**Attribute levels**	**β (mean)**	**SE**	**β (mean)**	**SE**	**IS (rank)**	**β (mean)**	**SE**	**IS (rank)**
ASC	-	0.13***	0.04	0.06*	0.03	-	0.13***	0.04	-
Choice of care providers	50% (ref)	-2.90***	0.14	-2.90***	0.13	0.62 (1)	-2.84***	0.15	0.55 (1)
	80%	0.20***	0.04	0.15***	0.04		0.21***	0.04	
	100%	2.70***	0.14	2.75***	0.13		2.64***	0.15	
Primary focus of provider contracting	Costs (ref)	-0.32***	0.04	-0.21***	0.03	0.06 (4)	-0.31***	0.04	0.08 (4)
	Quality	0.44***	0.04	0.34***	0.04		0.44***	0.04	
	SR	-0.12***	0.04	-0.13***	0.03		-0.13***	0.04	
Level of service benefits	Negative (ref)	-0.66***	0.04	-0.59***	0.03	0.13 (3)	-0.67***	0.04	0.13 (3)
	Positive	0.66***	0.04	0.59***	0.03		0.67***	0.04	
Premium per month	Per €	-0.05***	<0.01	-0.04***	<0.01	0.20 (2)	-0.05***	<0.01	0.43 (2)
Attributes	**Attribute levels**	**β (SD)**	**SE**	**β (SD)**	**SE**		**β (SD)**	**SE**	
ASC	-	0.01	0.07	<0.01	0.10		0.01	0.09	
Choice of care providers	80%	0.49***	0.03	0.55***	0.04		0.43***	0.05	
	100%	2.08***	0.08	2.23***	0.11		1.92***	0.12	
Primary focus of provider contracting	Quality	0.37***	0.03	0.38***	0.04		0.37***	0.05	
	SR	0.03	0.07	<0.01	0.12		0.09	0.10	
Level of service benefits	Positive	0.48***	0.02	0.47***	0.03		0.51***	0.04	
Interactions		**β (mean)**	**SE**						
Mode * ASC	-	-0.07	0.05	-	-	-	-	-	-
Mode * choice	80%	-0.05	0.06	-	-	-	-	-	-
	100%	0.01	0.17	-	-	-	-	-	-
Mode * contracting	Quality	-0.11**	0.05	-	-	-	-	-	-
	SR	<0.01	0.05	-	-	-	-	-	-
Mode * service benefits	Positive	-0.08*	0.05	-	-	-	-	-	-
Mode * premium	Per €	0.01***	<0.01	-	-	-	-	-	-
Model fit									
Number of observations		13449		7995			5454		
AIC		0.95		0.94			0.96		
Pseudo-R2		0.32		0.32			0.31		
Log likelihood		-6356.84		-3750.12			-2602.07		

Abbreviations: *SE* standard error; *IS* relative importance; *Ref* reference level; *SD* standard deviation; *SR* social responsibility; *AIC* Akaike Information Criterion.
Notes: [1]*, **, *** denote significance at the 0.1, 0.05 and 0.01 level respectively; [2]The attributes 'choice of care providers', 'primary focus of provider contracting', and 'level of service benefits' were effects coded. Premium was coded as a linear attribute; [3]Random parameters are assumed to be normally distributed; [4]Number of Halton draws used to estimate the model: 1000; [5]Mode was coded as 1 for online respondents, and 0 for paper respondents.

Table 5 | Average marginal WTP estimates in euros per month.

		Online sample (n=533)		Paper sample (n=365)		Ratio
Attributes	**Attribute levels**	**WTP (€)**	**95% CI[1]**	**WTP (€)**	**95% CI[1]**	
Choice of care providers	50 → 80%	76	69;84	58	52;63	1.32
	50 → 100%	142	125;159	104	90;117	1.36
Primary focus of provider contracting	Price → quality	14	13;14	14	13;15	0.98
	Price → SR	2	0.88;4	3	2;5	0.63
Level of service benefits	Negative → positive	29	27;32	25	23;27	1.16

Abbreviations: *WTP* willingness-to-pay; *CI* confidence interval; *SR* social responsibility.
Notes: [1]95% confidence intervals were approximated using the Delta method [22].

DISCUSSION

The median response time was shorter for the online sample, compared to the paper sample, and a smaller proportion of respondents from the online sample was satisfied with the number of choice sets. In addition, a larger proportion of respondents from the online sample, compared to the paper sample, failed the monotonicity axiom. Although we found that some WTP estimates were higher for the online sample, the elicited preferences for basic health insurance characteristics were similar between both modes of administration.

In previous studies, Windle & Rolfe [24] found that Internet respondents were more confident that they made the correct choices and that they found the choice sets less confusing. Olsen [25] showed that although Internet respondents reported to be more certain of their choices, the estimation precision and reliability in choices was higher in the mail sample. Findings of these studies, as well as our findings could be explained by the fact that Internet panel members are experienced in completing online surveys, possibly resulting in less well considered choices. Savage et al. [26] compared respondents' learning and fatigue between mail and online DCE administration. They found that online respondents who completed the DCE suffered from fatigue or boredom, while mail respondents did not and where therefore more consistent in their answers [26]. We found some significant between-sample differences in choice probabilities for the last five choice sets, which supports the recommendation of Savage et al. that researchers might need to include fewer choice sets per respondents when collecting DCE data online, compared to when collecting DCE data using paper surveys. In order to collect the same number of choice observations, researchers might want to increase the number of respondents that participate in the online DCE. This may be feasible since the costs for online data-collection are usually lower. The price per completed online survey in our study was €4 excluding taxes, while the price per completed paper survey was €14. Previous studies also found that conducting a DCE on paper was more expensive than conducting a DCE online [24, 25]. For instance, Windle & Rolfe reported $70 per paper survey and $15 per online survey [24]. In contrast to previous

studies in which lower (average ratio 0.76) [12] and similar WTP estimates were found [24, 25] for online responders versus paper responders, we found some of the WTP estimates in our study to be slightly higher for online respondents compared to paper respondents, while other WTP estimates were equal between the samples.

This study is subject to limitations. First, as opposed to previous studies with a similar aim, our findings are less prone to bias stemming from differences in respondent characteristics between samples. We successfully controlled for well-known confounders that may distort the comparison of paper survey and online responses [27, 28] and indeed found no significant differences in respondent characteristics. Unfortunately, we were not able to randomize respondents to either one of the administration modes, nor did we use a crossover design in which respondents answered the choice sets both online and on paper, and there might therefore be differences in unobserved respondent characteristics that could have biased our results. Second, since we aimed to study measurement effects, not sample composition effects, we conducted our study using the same sample frame for both administration modes. The sample frame we used might have affected our results because Internet panel members are experienced to completing surveys online, not on paper. In addition, these Internet panel members have actively signed-up to complete surveys. Results might have shown more differences between administration modes if we would have used other samples of respondents, e.g. random digit dialling versus an Internet panel. Third, we used a DCE on health insurance preferences as a case for this methodological study. This DCE was administered among a sample representative of the general adult population of the Netherlands. The generalizability of our findings to other respondent groups (e.g. elderly) or other topics may be limited.

Our study also provides directions for future research. Due to its technical possibilities, electronic data-collection offers additional benefits over paper data-collection. For example, explanations using video might improve the survey understanding of lower literate respondents [29]. It would be worthwhile if future DCE studies test whether such features aid in the understanding of the DCE. Electronic data-collection also provides more control and flexibility over the information that is presented in a survey [30] compared to paper data-collection and can therefore more easily facilitate methodologic DCE studies that incorporate split samples with different presentation formats, e.g. when studying ordering effects [10]. Previous research shows no mode effects between a mobile app and computer web survey administration [31]. Although we did not receive any negative feedback hereof in the current study, it can be questioned whether the choice sets of a DCE are fully visible on the relatively small smartphone screens. Future studies might therefore want to focus on the effects of completing a DCE on different digital devices, such as laptops, tablets, and smartphones.

In conclusion, we find no indication that online surveys yield inferior results compared to paper based surveys, while the price per respondent is lower for online surveys. However, researchers might want to include fewer choice sets per respondent when collecting DCE data online. More research is needed to support our findings.

REFERENCES

1. Bridges JF, Hauber AB, Marshall D, et al. Conjoint analysis applications in health--a checklist: a report of the ISPOR Good Research Practices for Conjoint Analysis Task Force. Value Health. 2011;14:403-13.
2. Clark MD, Determann D, Petrou S, et al. Discrete Choice Experiments in Health Economics: A Review of the Literature. Pharmacoeconomics. 2014;32:883-902.
3. Ekman A, Litton JE. New times, new needs; e-epidemiology. Eur J Epidemiol. 2007;22:285-92.
4. Van Gelder MM, Bretveld RW, Roeleveld N. Web-based questionnaires: the future in epidemiology? Am J Epidemiol. 2010;172:1292-8.
5. Dillman DA. Why choice of survey mode makes a difference. Public Health Rep. 2006;121:11-3.
6. Bowling A. Mode of questionnaire administration can have serious effects on data quality. J Public Health. 2005;27:281-91.
7. Lindhjem H, Navrud S. Using internet in stated preference surveys: a review and comparison of survey modes. Int Rev Environ Resour Econ. 2011;5:309-51.
8. Jäckle A, Roberts C, Lynn P. Assessing the Effect of Data Collection Mode on Measurement. Int Stat Rev. 2008;78:3-20.
9. Howard K, Salkeld G. Does attribute framing in discrete choice experiments influence willingness to pay? Results from a discrete choice experiment in screening for colorectal cancer. Value Health. 2009;12:354-63.
10. Kjaer T, Bech M, Gyrd-Hansen D, et al. Ordering effect and price sensitivity in discrete choice experiments: need we worry? Health Econ. 2006;15:1217-28.
11. Veldwijk J, Lambooij MS, van Til JA, et al. Words or graphics to present a Discrete Choice Experiment: Does it matter? Patient Educ Couns. 2015;98:1376-84.
12. Boyle KJ, Morrison M, MacDonald DH, et al. Investigating Internet and Mail Implementation of Stated-Preference Surveys While Controlling for Differences in Sample Frames. Environ Resource Econ. 2015:1-19.
13. Pieterse AH, Berkers F, Baas-Thijssen MC, et al. Adaptive Conjoint Analysis as individual preference assessment tool: feasibility through the internet and reliability of preferences. Patient Educ Couns. 2010;78:224-33.
14. Mulhern B, Longworth L, Brazier J, et al. Binary choice health state valuation and mode of administration: head-to-head comparison of online and CAPI. Value Health. 2013;16:104-13.
15. Cairns J, van der Pol M. Repeated follow-up as a method for reducing non-trading behaviour in discrete choice experiments. Soc Sci Med. 2004;58:2211-8.
16. Lancsar E, Louviere J. Deleting 'irrational' responses from discrete choice experiments: a case of investigating or imposing preferences? Health Econ. 2006;15:797-811.
17. Lloyd AJ. Threats to the estimation of benefit: are preference elicitation methods accurate? Health Econ. 2003;12:393-402.
18. Malhotra N. Completion time and response order effects in web surveys. Public Opin Quart. 2008;72:914-34.
19. Orme BK. Getting started with Conjoint Analysis: Strategies for Product Design and Pricing Research. Madison, WI: Reserach Publishers LLC; 2006.
20. De Bekker-Grob EW, Donkers B, Jonker MF, et al. Sample Size Requirements for Discrete-Choice Experiments in Healthcare: a Practical Guide. Patient. 2015;8:373-84.
21. Greene WH, Hensher DA. A latent class model for discrete choice analysis: contrasts with mixed logit. Transport Res B-Meth. 2003;37:681-98.
22. Hole AR. A comparison of approaches to estimating confidence intervals for willingness to pay measures. Health Econ. 2007;16:827-40.
23. Fransen MP, Van Schaik TM, Twickler TB, et al. Applicability of internationally available health literacy measures in the Netherlands. J Health Comm. 2011;16Suppl3:134-49.

24. Windle J, Rolfe J. Comparing responses from internet and paper-based collection methods in more complex stated preference environmental valuation surveys. Econom Anal Policy. 2011;41:83-97.
25. Olsen SB. Choosing between internet and mail survey modes for choice experiment surveys considering non-market goods. Environ and Resour Econ. 2009;44:591-610.
26. Savage SJ, Waldman DM. Learning and fatigue during choice experiments: a comparison of online and mail survey modes. J Applied Econ. 2008;23:351-71.
27. Barentsz MW, Wessels H, van Diest PJ, et al. Tablet, Web-Based, or Paper Questionnaires for Measuring Anxiety in Patients Suspected of Breast Cancer: Patients' Preferences and Quality of Collected Data. J Med Internet Res. 2014;16:e239.
28. Diment K, Garrett-Jones S. How demographic characteristics affect mode preference in a postal/web mixed-mode survey of Australian researchers. Soc Sci Comput Rev. 2007;25:410-7.
29. Wright KB. Researching Internet-Based Populations: Advantages and Disadvantages of Online Survey Research, Online Questionnaire Authoring Software Packages, and Web Survey Services. J Comput-Mediat Commun. 2005;10:00-00.
30. Tingling P, Parent M, Wade M. Extending the capabilities of Internet-based research: lessons from the field. Internet Research. 2003;13:223-35.
31. Wells T, Bailey JT, Link MW. Comparison of smartphone and online computer survey administration. Soc Sci Comput Rev. 2014;32:238-55.

10

Exploring how individuals complete the choice tasks in a discrete choice experiment: an interview study

*Jorien Veldwijk
*Domino Determann
Mattijs S. Lambooij
Janine A. van Til
Ida J. Korfage
Esther W. de Bekker-Grob
G. Ardine de Wit

BMC Medical Research Methodology. 2016;16:45
doi: 10.1186/s12874-016-0140-4

* Authors contributed equally to this manuscript

ABSTRACT

Background: To be able to make valid inferences on stated preference data from a Discrete Choice Experiment (DCE) it is essential that researchers know if participants were actively involved, understood and interpreted the provided information correctly and whether they used complex decision strategies to make their choices and thereby acted in accordance with the continuity axiom.

Methods: During structured interviews, we explored how 70 participants evaluated and completed four discrete choice tasks aloud. Hereafter, additional questions were asked to further explore if participants understood the information that was provided to them and whether they used complex decision strategies (continuity axiom) when making their choices. Two existing DCE questionnaires on rotavirus vaccination and prostate cancer-screening served as case studies.

Results: A large proportion of the participants was not able to repeat the exact definition of the risk attributes as explained to them in the introduction of the questionnaire. The majority of the participants preferred more optimal over less optimal risk attribute levels. Most participants (66%) mentioned three or more attributes when motivating their decisions, thereby acting in accordance with the continuity axiom. However, 16 out of 70 participants continuously mentioned less than three attributes when motivating their decision. Lower educated and less literate participants tended to mention less than three attributes when motivating their decision and used trading off between attributes less often as a decision-making strategy.

Conclusion: The majority of the participants seemed to have understood the provided information about the choice tasks, the attributes, and the levels. They used complex decision strategies (continuity axiom) and are therefore capable to adequately complete a DCE. However, based on the participants' age, educational level and health literacy additional, actions should be undertaken to ensure that participants understand the choice tasks and complete the DCE as presumed.

INTRODUCTION

A Discrete Choice Experiment (DCE) is a stated preference method in which individuals are asked to choose between two or more scenarios. Each scenario consists of several attributes with systematically varying levels that describe the product or service at hand. By monitoring individuals' choices over a series of choice tasks, their preferences are elicited. DCEs are increasingly being used to make inferences on individuals' preferences for a wide range of products or services within a health care context [1, 2].

DCE results are analysed according to economic theory like Lancaster's theory of demand [3], random utility theory [4, 5] or random regret minimization [6, 7]. These methodologies use a multi-attribute approach [8, 9]. It might be that individuals do not understand all the information that was provided to them and do not weigh all attributes when making their choices, especially not if risk information is included [10-12]. Therefore, this methodological approach may result in invalid conclusions regarding the attribute level estimates and the estimated potential uptake rates of goods or services. This in turn may lead to sub-optimal concordance between stated and revealed preferences. For these reasons, it is essential that researchers know how participants interpret the attributes and the levels, and ultimately make their decision.

Theoretical assumptions

Conducting, analysing and interpreting DCEs is based on several implicit and explicit assumptions regarding respondents' decision-making, among which the ones listed next [13-15]. It is assumed that respondents are actively involved in completing the choice tasks. Additionally, respondents are expected to understand and interpret the information that they are provided with, as intended by the researcher [16, 17]. Finally, respondents are assumed to use complex decision strategies by considering all attributes and making their choice based on trade-offs between all attributes (continuity axiom) [7, 18, 19].

Theoretical assumptions in practice

Both within and outside the health care setting, mainly quantitative research showed that these assumptions do not always hold. First, health-related DCEs often contain risk attributes. Research showed that respondents often misinterpret risk information [20, 21]. For example, respondents often interpret numerical values (ratio scales) as categorical information (for instance respondents recode a risk of 10%, 30% or 50% to a low-medium-high risk) in DCEs [22-24]. In addition, respondents might apply simplified decision strategies such as choosing a scenario based on one attribute only [25]. Such simplifying strategies may especially be used by lower educated and less health literate respondents [8]. Second, completing choice tasks can be a cognitive challenge [26, 27]. Cognitively demanding decisions induce the use of simplified heuristics [28-31], which is not in accordance with the assumption that people

use complex decision strategies; hence, people do not act in accordance with the continuity axiom. Additional research on this latter axiom showed that participants with dominant preferences base their decisions on one high priority attribute [32]. Such non-compensatory decision-making could either reflect a true strong preference for one specific attribute or it may be a way to avoid complex decision-making [33, 34]. Moreover, different quantitative studies show that up to 45% of the participants have dominant preferences [33, 35, 36] and that lower educated participants more often base their decisions on dominant preferences [33]. Other studies showed that participants may disregard certain attributes and base their decision on some, but not on all attributes (attribute-non-attendance) [24, 32, 34, 37-42], thereby violating the continuity axiom.

Aims

This study explored in depth how respondents complete choice tasks in a DCE, whether participants were actively involved, understood and interpreted the provided information correctly and whether they used complex decision strategies to make their choices and thereby acted in accordance with the continuity axiom. It was tested whether results differed by respondents' educational level and health literacy. In contrast to other published qualitative studies that used a retrospective 'top-down' approach in relatively small samples to determine if and why respondents violate theoretical axioms, this paper uses a prospective 'bottom-up' approach in a large sample, and specifically focusses on respondents' understanding and interpretation of risk information, and their use of complex decision-making strategies.

MATERIALS AND METHODS

Discrete Choice Experiments

Two previously administered Dutch DCE questionnaires, that used a state-of-the-art approach by designing their experiment according to the latest guidelines for DCEs [13, 15], were used as case studies for the current study [43, 44]. One DCE reported on parental preferences for rotavirus vaccination while the other DCE reported on men's preferences for prostate cancer-screening. Both DCEs selected their attributes and designed the survey based on formal literature review, interviews with experts, focus group discussions with participants, a pilot study and a think-aloud pilot study. Additionally, both DCEs contained several risk attributes, namely: vaccine effectiveness & frequency of severe side effects (rotavirus DCE) and proportion of unnecessary biopsies & proportion of unnecessary treatment (prostate cancer-screening DCE). Detailed descriptions of both studies are reported elsewhere [43, 44]. Since DCEs often cover very specific health topics, and thereby have very selective study samples, we included two DCEs to increase participant heterogeneity regarding demographic characteristics. A sample of the respondents of the case studies was re-contacted after

previously indicating that they were willing to participate in further research. Participants completed the initial DCE at least 6 months before the interview. See additional file 1 for a description of both studies, Table 1 and 2 for a description of the included attributes and levels, and additional file 2 for examples of choice tasks of both case studies.

Table 1 | Attributes and levels for rotavirus DCE.

Attributes	Explanation	Levels
Vaccine effectiveness	The percentage of children that will be protected against a rotavirus infection when vaccinated	55%
		75%
		95%
Frequency of severe side effects	The number of vaccinated children that will suffer from intussusception due to vaccination[1]	1 in 10,000
		1 in 100,000
		1 in 1,000,000
Protection duration	The number of years that the vaccine protects against a rotavirus infection [47]	1 year
		3 years
		6 years
Healthcare facility of vaccine administration		Child welfare centre
		General practitioner
Out-of-pocket costs		€0
		€30
		€140

Note: [1]Intussusception is an acute condition in which part of the bowel telescopes into another adjacent part of the bowel, resulting in obstruction.

Participants

In total, we included 70 participants for the current study; 35 from the rotavirus DCE and 35 from the prostate cancer-screening DCE. To study potential differences in decision-making strategies between lower and higher educated respondents, we purposively sampled equal proportions of lower and higher educated individuals from the participants of the previously performed DCE's who had indicated to be willing to participate in future research. If subjects agreed to participate in the current study, they received a package with materials by mail. The Dutch National Ethics Board (Central Committee on Research involving Human Subjects) concluded that formal testing by a medical ethical committee was not necessary as participants only completed one non-invasive questionnaire on voluntary basis. Results were not analysed or reported at the individual level, which is in accordance with the guidelines laid down in the Declaration of Helsinki.

Table 2 | Attributes and levels for prostate cancer-screening DCE.

Attributes	Explanation	Levels
Number of deaths from prostate cancer	It was given that 35 out of 1000 men die because of prostate cancer when no screening program is provided	32 deaths (3 deaths prevented)
		28 deaths (7 deaths prevented)
		25 deaths (10 deaths prevented)
		18 deaths (17 deaths prevented)
Frequency of blood test		Every year
		Every 2 years
		Every 3 years
		Every 4 years
Number of unnecessary biopsies	Number of men, per 1000 men with an elevated PSA level, in which biopsies are unnecessary. Unnecessary biopsies were defined as biopsies in which no cancer was found, but in which PSA levels suggested that there was cancer	200 unnecessary biopsies (800 justified biopsies)
		400 unnecessary biopsies (600 justified biopsies)
		600 unnecessary biopsies (400 justified biopsies)
		800 unnecessary biopsies (200 justified biopsies)
Number of unnecessary treatments	Number of men, per 1000 treated men, in whom treatment is unnecessary. Unnecessary treatment was defined as treatment that was not life prolonging, however it could lead to urine-loss and erection disorders due to treatment	0 unnecessary treatments (1000 justified treatments)
		200 unnecessary treatments (800 justified treatments)
		500 unnecessary treatments (500 justified treatments)
		800 unnecessary treatments (200 justified treatments)
Out-of-pocket costs per year		€0
		€50
		€100
		€300

Interviews

Both face-to-face (N=5 per cohort) and telephone interviews (N=30 per cohort) were scheduled. Interview guides were developed for both DCEs. During a consensus meeting with all authors the categorization of answers was discussed. Although the topic of the two DCEs differed, both guides described a similar interview protocol to make the results of both groups comparable. The structured interviews were pilot tested (N=7) to optimize the interview guide, to test the duration of an interview and to ensure both interviewers conducted the interviews in the same manner. This resulted in minimal adaptations to the interview guides. The final interview outline is described in Table 3. All interviews started with a short introduction to the current study. Next, participants were given some time to read the introduction of the DCE questionnaire. To get familiar with the DCE and the think aloud method, participants were asked to complete one choice task as a warm up exercise. The core of the interview consisted of three parts. During part one (think aloud part), participants completed four choice tasks from the original DCE (Table 3). We instructed the participants to think aloud when reading and completing the choice tasks. Part one of the interview took place without any specific guidance by the interviewers in order to mimic non-lab questionnaire completion situations as much as possible. However, if a participant was quiet for some time, the interviewer reminded him/her to keep thinking aloud and to report his/her thoughts. During part two of the interview (interview part), specific questions were asked to test the interpretation of the risk attributes, the understanding of the risk attributes, the decision strategy and the continuity axiom (Table 3). Finally, in part three of the interview, health literacy was measured both by means of a subjective self-reported questionnaire [45] and a validated objective measurement [45] (see additional file 3). Results will be reported in the following order: choice task reading, interpretation of the risk attributes, understanding of the risk attributes, decision strategy and continuity axiom and differences by educational level and health literacy (stratified by the rotavirus and the prostate cancer-screening cohort).

Two researchers (JV and DD) conducted the interviews. The interviewers used a predefined form to categorize reading and decision-making behaviour in part one (this for instance entailed monitoring and marking how individuals read the choice tasks), as well as the answers the participants provided in part two and three of the interview (see Table 3). They also made notes and wrote down specific observations during each interview. Interviews were audio taped. Whenever there was doubt about participants' behaviour (in part one) or their answers (in part two), the two interviewers discussed and jointly listened to the audiotaped interview and completed the predefined form. As a result of this use of objective and pre-specified categories in the interviews, data could be analysed with SPSS.

Table 3 | Interview outline.

Short introduction to the current study including a choice task as a warm up exercise to get used to the DCE and thinking aloud		
Part 1: Think aloud part (categorization of participants observed decision-making behaviour over four choice tasks, no specific questions asked)		
		Categorization options
Choice task reading	In which manner participants read the choice tasks	Attribute-wise
		Scenario-wise
		Directly motivating decision
		Otherwise
	Whether participants from the prostate cancer-screening cohort read the opt-out option aloud	Yes
		No
Interpretation of the risk attributes	How participants mentioned the risk attributes	Mentioning actual values
		Translating levels into ordinal scale
		Mentioning and interpreting values
Testing of continuity axiom	The number of attributes participants mentioned when motivating their decision for a certain scenario. Participants were marked as acting in accordance with the continuity axiom if they mentioned three or more attributes (i.e. less than the majority of the five included attributes) when motivating their decision	One
		Two
		Three or more
Decision strategy	The decision strategies participants applied to make their decision	Traded off attribute levels
		Based decision on one attribute
		Otherwise

Part 2: Interview part (asking direct questions)		
	Questions asked	**Answer categories**
Interpretation of the risk attributes	[1] was one of the characteristics that was included in the choice tasks. What did you have in mind with respect to this characteristic when you completed the choice tasks?	Exact definition Other definition
Understanding of the risk attributes	Please look at choice task x. If you were asked to make a choice based on [1] only, which scenario would you choose? This question was asked twice for all tested risk attributes (see[1]).	Scenario 1 Scenario 2 Don't know
	Control question: Participants were asked to make a simple calculation with respect to the risk attributes to test their understanding of the numerical values of the risk attributes. For the rotavirus cohort: 'Imagine, 1.000 children will get vaccinated with a vaccine that is 95% effective. Assume that all children will get in contact with the virus. How many children will not get sick?', and 'Imagine, 300.000 children will get the rotavirus vaccine. Assume that the vaccine will lead to severe side effects in 1 out of every 100.000 children. How many children will suffer from severe side effects? For the prostate cancer-screening cohort: 'Imagine a screening program in which out of 1.000 treatments, 200 are unnecessary. Imagine that 2.000 men participate in this screening program. How many men will be treated unnecessarily?'	Right answer Wrong answer Don't know
Testing of continuity axiom	Those participants that based their decision on less than three attributes in all choice tasks were asked: 'You included only x out of five characteristics when making your choice. Why was this the case?	Only one or two attributes important Hard to trade off multiple attributes Lack of attribute understanding
Part 3: Measuring health literacy[2]		
Subjective health literacy	Set of Brief Screening Questions (SBSQ-D) of Chew for prostate cancer-screening cohort only. This instrument was already included in the initial rotavirus DCE and was therefore not repeated in the current study.	
Objective health literacy	Newest Vital Sign (NVS-D)	

Notes: [1]For the rotavirus cohort, these questions were asked for both the attributes vaccine effectiveness and frequency of severe side effects, while for the prostate cancer-screening cohort, these questions were only asked for the unnecessary treatment attribute since the levels of the two selected risk attributes (unnecessary treatment and unnecessary biopsy) were considered to be equal. [2]See additional file 3 for more information on these instruments.

RESULTS

Table 4 describes the demographic characteristics of the participants who were interviewed. The average duration of the interviews in the rotavirus cohort was 27 minutes, while the average duration of the interviews in the prostate cancer-screening cohort was 41 minutes.

Table 4 | Demographics of participants in both cohorts.

		Rotavirus cohort (n=35)	Prostate cancer-screening cohort (n=35)
		Mean (SD)	**Mean (SD)**
Age in years		30.4 (4.5)	67.6 (5.5)
		Proportion (%)	**Proportion (%)**
Gender	Female	94.3	0
Education[1]	Lower	45.7	48.6
	Higher	54.3	51.4
Health literacy[2]	High subjective score	100	100
	High objective score	100	55.9

Notes: [1]Educational level was dichotomized into a higher and a lower educational level, whereby a Bachelor's and/or Master's degree were defined as a higher educational level and all other educational levels were defined as a lower educational level. [2]High subjective score includes participants with a score >2 on the SBSQ-D. High objective score includes participants with a score of 4-6 on the NVS-D.

Choice task reading

a. Think aloud part

Within both cohorts, the majority of the participants (60.7% for the rotavirus cohort, and 56.4% for the prostate cancer-screening cohort) read the choice tasks attribute-wise, starting from the top and moving to the bottom. In the rotavirus cohort, two other frequently used strategies for reading the choice tasks were 1) reading scenario-wise (15.0%), and 2) directly motivating which of the two scenarios was preferred based on the attribute levels (14.3%). This latter strategy was also often applied in the prostate cancer-screening cohort (18.6%). Additionally, a considerable number of participants used different reading strategies (12.1%); only reading attributes that were of personal importance, only reading attributes that differed between the two scenarios, and reading choice tasks (completely) in a random manner. The prostate cancer-screening choice tasks included an opt-out option (i.e. no screening), that was specifically read aloud by 42.9% of the participants in choice task one, by 25.7% in choice task two, 20.0% in choice task three and 8.6% of the participants in choice task four.

Interpretation of the risk attributes

a. Think aloud part

With respect to the risk attributes, on average over all four choice tasks, 56.6% of the participants of the rotavirus cohort mentioned the actual values of the attribute levels

for vaccine effectiveness while completing the choice task and 45.9% mentioned this for frequency of severe side effects. For the attribute 'vaccine effectiveness', on average over the four choice tasks, 17.5% of the participants described the levels on an ordinal scale and 20.6% combined reading with interpretation, like: *'In total 75 out of every 100 children are protected against a rotavirus infection, or three-quarters of the children do not become ill'.* With respect to the frequency of severe side effects, these percentages were 23.7% and 20.6% respectively.

In the prostate cancer-screening cohort, 52.2% of the participants mentioned the actual values of both of these attributes when reading the choice tasks. Additionally, 12.9% of the participants interpreted the number of unnecessary biopsies and 14.3% of the participants interpreted the number of unnecessary treatments when reading the choice tasks, for example: *'If I have to choose between 200 or 800 unnecessary biopsies/treatments, the likelihood of me having an unnecessary biopsies/treatment is four times as high in scenario two'.* Others did not mention these attributes while reading the choice tasks (30.7% for the number of unnecessary biopsies and 29.3% for the number of unnecessary treatments). Many of the participants experienced difficulties interpreting these two attributes. Some participants who experienced such difficulties did not understand the difference between biopsies and treatment, and some even thought they were similar or at least had similar side effects. For instance, participants stated: *'An unnecessary biopsy is an unnecessary treatment'* or *'Biopsy causes urine incontinence'.* Some participants stated that they ignored these attributes when reading the choice tasks for those reasons, while others misinterpreted the numbers.

b. Interview part

Twenty percent of the participants of the rotavirus cohort was able to repeat the definition of vaccine effectiveness as described in the introduction section of the questionnaire. Another 57.1% described vaccine effectiveness as 'how well a vaccine works' and 22.9% provided a completely different definition. When asked about the meaning of the attribute side effects, the definition of side effects as provided in the questionnaire was mentioned by 37.1% of the participants, 54.1% interpreted side effects correctly but mentioned additional side effects that were not mentioned in the explanation of the attribute, such as a high temperature, feeling sick or dying, while 11.4% provided a completely different definition.

In the prostate cancer-screening cohort, only 17.1% of the participants was able to give the definition of the unnecessary treatment attribute as described in the attribute explanation section of the questionnaire.

Understanding of the risk attributes

a. Interview part

All participants of the rotavirus cohort chose the vaccine with the highest effectiveness within both choice tasks when they were asked to choose based on this one attribute. On average over two choice tasks, all but three (4.3%) participants chose the scenario with the lowest frequency of severe side effects. 77.1% of the participants gave the correct answer to the control question for vaccine effectiveness, and 94.3% of the participants gave the right answer to the control question for frequency of severe side effects. These results indicate that most participants were able to interpret percentages and frequencies correctly.

Within the prostate cancer-screening cohort, 83% chose the screening option with the lowest level of unnecessary treatments. Although the concepts might not have been completely clear to some participants, 88.6% answered the control question correctly, indicating that the participants were able to interpret the numbers of unnecessary treatment correctly.

Decision strategy and continuity axiom

a. Think aloud part

In both cohorts, most participants mentioned the majority of the included attributes while motivating their choice for a scenario, which is in accordance with the continuity axiom (Table 5). In both cohorts, the majority also traded off between the levels of those attributes when motivating their decision, which again is in accordance with the continuity axiom. Within the rotavirus cohort, 20.0% mentioned two attributes and 7.2% only mentioned one attribute when motivating their decisions. In the prostate cancer-screening cohort 16.4% mentioned two attributes, 17.9% only mentioned one attribute and 5.7% did not mention any of the attributes but chose to opt-out.

b. Interview part

A total number of 16 participants (one in the rotavirus cohort and 15 in the prostate cancer-screening cohort) continuously traded off less than three attributes when completing the choice tasks. Nine out of those 16 participants stated that they traded off so few attributes because only those attributes were important to them, the other seven mentioned that they did so because they found it hard to trade off more attributes at once or because they did not understand the meaning of certain attributes. This latter category of seven participants comprised of participants for whom it is questionable whether they grasped the questions and understood the hypothetical nature of the choice tasks at all. The finding that some participants might not have understood the DCE at all is reflected in the fact that they decided per attribute which scenario they preferred, without making one final decision for one scenario. They also mentioned things such as: *'What is the difference between this question and the previous one?'* or *'Can I switch between scenarios within one question?'*

Table 5 | Continuity axiom and decision strategy.

		Average over all four choice tasks (%)
Rotavirus cohort (n=35)	**Motivating decision (continuity axiom)[1]**	
	Motivation based on one attribute	7.2
	Motivation based on two attributes	20.0
	Motivation based on three or more attributes	72.9
	Decision strategy for those who acted in accordance with the continuity axiom	
	Traded off attribute levels between each other	85.6
	One attribute was most decisive	11.5
	Otherwise	2.9
Prostate cancer-screening cohort (n=35)	**Motivating decision (continuity axiom)[1,2]**	
	Motivation based on one attribute	17.9
	Motivation based on two attributes	16.4
	Motivation based on three or more attributes	60.0
	Decision strategy for those who acted in accordance with the continuity axiom	
	Traded off attribute levels between each other	60.0
	One attribute was most decisive	26.4
	Otherwise	13.6

Notes: [1]Participants were marked as acting in accordance with the continuity axiom, only if they motivated their decision based on three or more attributes. [2]These numbers do not add up to 100% because some men did not mention any of the attributes when motivating which scenario they preferred; they chose opt-out (5.7%).

Differences by educational level and health literacy

Overall, there is a trend showing that more educated and literate participants included three or more attributes when motivating their decision and that they traded off between attributes more often compared to participants with a lower educational level or lower health literacy score (Table 6). Additionally, higher educated and literate participants more often correctly explained the risk attributes and more often answered the risk attribute control question correctly (Table 6). Finally, lower educated and less literate participants who based their decision on two attributes or less, more often stated that they found it difficult to compare all attributes.

Table 6 | Differences in educational level and health literacy[1].

	Rotavirus cohort Educational level (n=35)[2]		Prostate cancer-screening cohort Educational level (n=35)[2]	
	Lower (%)	Higher (%)	Lower (%)	Higher (%)
Including three or more attributes when motivating decisions	81.3	100.0	70.6	83.3
Trading off attribute levels as a strategy to make a decision	56.3	73.7	35.3	44.4
Right explanation of vaccine effectiveness	12.5	26.3	-	-
Right explanation of severe side effects	56.3	94.7	-	-
Right explanation of unnecessary treatments	-	-	11.8	22.2
Right answer to control question on vaccine effectiveness	18.8	52.6	-	-
Right answer to control question on severe side effects	87.5	100.0	-	-
Right answer to control question on unnecessary treatments	-	-	82.4	94.4
			Health literacy (n=34)[3]	
			Low (%)	High (%)
Including three or more attributes when motivating decisions			80.0	73.7
Trading off attribute levels to make a decision			33.3	47.4
Right explanation of unnecessary treatments			6.7	21.1
Right answer to control question on unnecessary treatments			80.0	94.7
			Combined measure (n=20)[4]	
			Low (%)	High (%)
Including three or more attributes when motivating decisions			77.8	81.8
Trading off attribute levels to make a decision			33.3	54.5
Right explanation of unnecessary treatments			0.0	18.2
Right answer to control question on unnecessary treatments			77.8	100.0
Perceived it as difficult to trade off >2 attributes			60.0	33.3

Notes: [1]Differences in health literacy could only be calculated for the prostate cancer-screening cohort, because 100% of the participants in the rotavirus cohort had high objective health literacy scores. [2]Educational level was dichotomized into a higher and a lower educational level, whereby a Bachelor's and/or Master's degree were defined as a higher educational level and all other educational levels were defined as a lower educational level. [3]High subjective score includes participants with a score >2 on the SBSQ-D. High objective score includes participants with a score of 4-6 on the NVS-D. [4]Individuals that scored low on both educational level and objective health literacy (n=9) or scored high on both educational level and objective health literacy (n=11).

DISCUSSION

The majority of the participants preferred more optimal over less optimal attribute levels and answered the control question(s) regarding their understanding of the numerical values of the risk attributes correctly. At the same time, a large proportion of the participants was not able to repeat the exact definition of the risk attributes as explained to them in the introduction of the questionnaire. While the majority of the participants based their decision on three or more attributes by trading them against each other, which implies complex decision strategies and is in accordance with the continuity axiom, about a third of the participants used simplifying strategies such as basing their decision on less than three attributes.

Acting in contrast with the continuity axiom does not seem to be a problem per se. In real life, individuals might also not include all product characteristics when making their decision. However, within a DCE analysis, this may result in invalid conclusions regarding the attribute level estimates and estimated potential uptake rates of goods or services, since a multi-attribute approach is undertaken to analyse the data [8, 9]. This in turn may lead to sub-optimal concordance between stated and revealed preferences. This is also reflected by previous studies that indicated different DCE outcomes and significant influences on marginal rates of substitution depending on attribute-non-attendance being taken into account in DCE analyses [32, 39-41]. Previous research described that this non-compensatory decision-making behaviour might have different causes; participants might actually have dominant preferences, it might be that the attribute levels are too similar, or that the participants lack understanding of certain attribute levels [18]. This latter was shown in the current study. In the rotavirus cohort for instance, 54% of the participants mentioned that they had other and sometimes far more serious side effects in mind when completing the choice tasks. This will probably cause an overestimation of the relative importance of the side effects attribute, which affects the WTP estimate. Additionally, in the prostate cancer-screening cohort, a majority of the participants indicated that they did not understand one or more attributes (mostly the risk attributes). Studies state that a lack of understanding of certain attribute (levels) might be due to a lower educational level, older age and a lower health literacy [8, 21, 23, 45, 46]. The current study indeed showed that the number of attributes included in decision-making, decision strategy, interpretation of the risk attributes and understanding of the risk attributes differed between participants with different educational levels and health literacy scores. This might also be reflected by the fact that the mean interview duration of the less literate and older prostate cancer-screening group was almost 15 minutes longer compared to the rotavirus cohort. Besides educational level and health literacy scores, the topic of the DCEs and the included attributes and attribute levels may have added to the differences that were found between the two cohorts.

This study was subject to some limitations. Firstly, the two DCEs that we used as case studies for this study were quite complex, because each included two risk attributes. It is commonly known that the interpretation of such attributes is perceived as more difficult by participants than for instance qualitative attributes [20, 48]. Difficulties in interpreting attribute levels and making decisions might therefore be more pronounced in this study compared to DCEs that include no or less risk-related attributes. However, since most health-related decisions include risk information, the case studies used for this study may be representative for many DCEs within a healthcare context. Secondly, this study focused on participants' understanding of the provided information on risk attributes, their use of complex decision strategies and the continuity axiom. Other assumptions underlying the DCE methodology, namely the rationality assumption (which does not describe the psychological assumption of rationality, but merely represents the completeness and transitivity axioms) and the monotonicity axiom, were not tested. Thirdly, although this study used the well-recognized think aloud method for the interviews, additional methods such as eye-tracking might provide even more insight into how and what participants read. Such research could focus on visual attention sequences and underlying decision processes, as well as reading strategies regarding for instance the opt-out option. The current study showed a decrease in the percentage of respondents reading the opt-out option, which might reflect that participants assume this option to be fixed (attribute levels are not changing). Additionally, eye-tracking research will also provide insight in the potential discrepancy between the way participants complete a DCE with or without thinking aloud. Future research could incorporate such methods when investigating participants' behaviour when completing a DCE questionnaire. Fourthly, although efforts were made to mimic non-lab choice situations, the fact that the interviewers were present during DCE completion might have influenced how participants completed the choice tasks. Participants therefore might have been more committed to completing the DCE. As a result, we might have overestimated the number of participants that acts in accordance with the tested assumptions. Fifthly, the sample size of 70 is relatively large for an interview study, at the same time, this sample size is too small to draw any conclusions based on statistical testing. However, the trends in the findings and the agreement of the current findings with the existing literature related to educational level and health literacy (non-DCE studies) provide face validity for the current study results. Confirmation of our findings is needed, e.g. from new DCEs including (preferably objective) health literacy measurements as well as axiom testing questions in their study.

The results of our study indicate that respondents have difficulties understanding all the information that is provided to them, they do not always use complex decision strategies to make their choices and therefore do not always act in accordance to the continuity axiom. This was most prominent in respondents with a lower educational level, higher age and lower health literacy status. We therefore recommend to conduct DCE questionnaires among older and/or less health literate populations in, for instance, mini-labs, where participants

complete DCEs in the presence of a researcher. Researchers have the opportunity to explain how to complete a DCE, including the hypothetical nature of the questionnaire and to answer questions that arise during the completion of the questionnaire, e.g. concerning the attributes and attribute levels. This is important especially among older target populations as participants in the prostate cancer-screening cohort sometimes indicated that they had difficulties interpreting the questions (e.g., *'In real life, I have a blood test to check my PSA levels every year, so I can only choose a scenario with that frequency of blood testing'*). This is in line with the findings of previous studies [37, 38]. Moreover, when conducting online research, the understanding of attribute levels among participants with a lower educational level and/or health literacy can be enlarged by providing the option to include an explanation of the attributes by audio or other technical solutions, e.g. pop-ups when clicking on attributes or levels. In addition, the option to listen to the explanation again while completing the choice tasks could be offered. Another recommendation is that a thorough pilot testing phase is necessary while developing a DCE, which includes think aloud testing to a priori identify possible problematic issues with the completion of the questionnaire. Finally, age, educational level and health literacy should be standard measures to include in every DCE questionnaire as well as in the analysis of DCE data. Until options to correct DCE responses for possible differences in demographic characteristics become common practice, researchers should at least describe these measures in their population and explain the possible effects on the results retrieved.

CONCLUSION

In conclusion, the majority of the participants seemed to have understood the provided information about the choice tasks, the attributes, and the levels. They used complex decision strategies (continuity axiom) and are therefore capable to adequately complete a DCE. However, based on the participants' age, educational level and health literacy additional actions should be undertaken to ensure that participants understands the choice tasks and complete the DCE as presumed.

REFERENCES

1. Clark MD, Determann D, Petrou S, et al. Discrete Choice Experiments in Health Economics: A Review of the Literature. Pharmacoeconomics. 2014;32:883-02.
2. De Bekker-Grob EW, Ryan M, Gerard K. Discrete choice experiments in health economics: a review of the literature. Health Econ. 2012;21:145-72.
3. Lancaster KJ. A new approach to consumer theory. J Polit Econ. 1966;132-57.
4. McFadden D: Conditional Logit Analysis of Qualitative Choice Behavior. In: Zarembka P. Frontiers in Econometrics. New York: Academic Press; 1974.
5. McFadden D. The Choice Theory Approach to Market Research. Mark Sci. 1986;5:275-97.
6. De Bekker-Grob EW, Chorus CG. Random regret-based discrete-choice modelling: an application to healthcare. Pharmacoeconomics. 2013;31:623-34.
7. Lancsar E, Louviere J. Deleting 'irrational' responses from discrete choice experiments: a case of investigating or imposing preferences? Health Econ. 2006;15:797-811.
8. Kenny P, Hall J, Viney R, et al. Do participants understand a stated preference health survey? A qualitative approach to assessing validity. Int Technol Assess Health Care. 2003;19:664-81.
9. Hensher DA. How do respondents process stated choice experiments? Attribute consideration under varying information load. J Applied Econ. 2006;21:861-78.
10. Harrison M, Rigby D, Vass C, et al. Risk as an attribute in discrete choice experiments: a systematic review of the literature. Patient. 2014;7:151-70.
11. Veldwijk J, Lambooij MS, van Til JA, et al. Words or graphics to present a Discrete Choice Experiment: Does it matter? Patient Educ Couns. 2015;98:1376-84.
12. Veldwijk J, van der Heide I, Rademakers J, et al. Preferences for Vaccination: Does Health Literacy Make a Difference? Med Decis Making. 2015;35:948-58.
13. Bridges JF, Hauber AB, Marshall D, et al. Conjoint analysis applications in health--a checklist: a report of the ISPOR Good Research Practices for Conjoint Analysis Task Force. Value Health. 2011;14:403-13.
14. Reed Johnson F, Lancsar E, Marshall D, et al. Constructing experimental designs for discrete-choice experiments: report of the ISPOR Conjoint Analysis Experimental Design Good Research Practices Task Force. Value Health. 2013;16:3-13.
15. Lancsar E, Louviere J. Conducting discrete choice experiments to inform healthcare decision making: a user's guide. Pharmacoeconomics. 2008;26:661-77.
16. Hensher DA, Rose JM, Greene WH. Applied Choice Analysis: A Primer. New York: Cambridge University Press; 2005.
17. Ryan M, Gerard K, Amaya-Amaya M. Using Discrete Choice Experiments to Value Health and Health Care. Dordrecht: Springer; 2008.
18. Cairns J, van der Pol M. Repeated follow-up as a method for reducing non-trading behaviour in discrete choice experiments. Soc Sci Med. 2004;58:2211-8.
19. Thurstone LL. The Method of Paired Comparisons for Social Values. J Abnorm Soc Psychol. 1927;21:384-400.
20. Galesic M, Garcia-Retamero R. Statistical numeracy for health: a cross-cultural comparison with probabilistic national samples. Arch Int Med. 2010;170:462-8.
21. Waters EA, Weinstein ND, Colditz GA, et al. Formats for improving risk communication in medical tradeoff decisions. J Health Commun. 2006;11:167-82.
22. Reyna VF, Brainerd CJ. Fuzzy-trace theory and framing effects in choice: Gist extraction, truncation, and conversion. J Behav Decis Making. 1991;4:249-62.
23. Lloyd AJ. Threats to the estimation of benefit: are preference elicitation methods accurate? Health Econ. 2003;12:393-402.
24. Ryan M, Watson V, Entwistle V. Rationalising the 'irrational': a think aloud study of discrete choice experiment responses. Health Econ. 2009;18:321-36.

25. Tversky A, Kahneman D. Judgment under uncertainty: Heuristics and biases. Science. 1974;185:1124-31.
26. De Bekker-Grob EW, Rose JM, Bliemer MC. A closer look at decision and analyst error by including nonlinearities in discrete choice models: implications on willingness-to-pay estimates derived from discrete choice data in healthcare. Pharmacoeconomics. 2013;31:1169-83.
27. Bryan S, Dolan P. Discrete choice experiments in health economics. For better or for worse? Eur J Health Econ. 2004;5:199-202.
28. Luce MF. Choosing to avoid: coping with negative emotion-laden consumer decisions. J Consum Res. 1998;24:409-33.
29. Luce MF, Payne JW, Bettman JR. Emotional trade-off difficulty and choice. J Market Res. 1999;36:143-59.
30. Iyengar SS, Kamenica E. Choice overload and simplicity seeking. Working paper. 2007. Available from http://www.cbdr.cmu.edu/seminar/Emir2.pdf.
31. Ritov I, Baron J. Status quo and omission biases. J Risk Uncertainty. 1992;5:49-61.
32. Lagarde M. Investigating attribute non-attendance and its consequences in choice experiments with latent class models. Health Econ. 2013;22:554-67.
33. Scott A. Identifying and analysing dominant preferences in discrete choice experiments: an application in health care. J Econ Psychol. 2002;23:383-98.
34. Alemu MH, Mørkbak MR, Olsen SB, et al. Attending to the reasons for attribute non-attendance in choice experiments. Environ Resour Econ. 2013;54:333-59.
35. Ryan M, Bate A. Testing the assumptions of rationality, continuity and symmetry when applying discrete choice experiments in health care. Appl Econ Letters. 2001;8:59-63.
36. Bech M, Kjaer T, Lauridsen J. Does the number of choice sets matter? Results from a web survey applying a discrete choice experiment. Health Econ. 2011;20:273-86.
37. Miguel FS, Ryan M, Amaya-Amaya M. 'Irrational' stated preferences: a quantitative and qualitative investigation. Health Econ. 2005;14:307-22.
38. Cheraghi-Sohi S, Bower P, Mead N, et al. Making sense of patient priorities: applying discrete choice methods in primary care using 'think aloud' technique. Family Practice. 2007;24:276-82.
39. Erdem S. Campbell D, Hole AR, Accounting for Attribute-Level Non-Attendance in a Health Choice Experiment: Does it Matter? Health Econ. 2015;24:773-89.
40. Kravchenko A. Influence of rudimentary attribute non-attendance (ANA) on choice experiment parameter estimates and design efficiency: A Monte Carlo Simulation analysis. J Choice Model. 2014;11:57-68.
41. Hole AR, Kolstad JR, Gyrd-Hansen D. Inferred vs. stated attribute non-attendance in choice experiments: A study of doctors' prescription behaviour. J Econ Behav Orga. 2013;996:21-31.
42. De Bekker-Grob EW, Rose JM, Donkers B, et al. Men's preferences for prostate cancer screening: a discrete choice experiment. Br J Cancer. 2013;108:533-41.
43. Veldwijk J, Lambooij MS, Bruijning-Verhagen P, et al. Parental preferences for rotavirus vaccination in young children: a Discrete Choice Experiment. Vaccine. 2014;32:6277-83.
44. Fransen MP, Van Schaik TM, Twickler TB, et al. Applicability of internationally available health literacy measures in the Netherlands. J Health Commun. 2011;16Suppl3:134-49.
45. Comparative report of health literacy in eight EU member states. The European health literacy survey HLS-EU2013. HLS-EU Consortium, 2012. Available from: www.health-literacy.eu.
46. Ozdemir S, Mohamed AF, Reed Johnson F, et al. Who pays attention in stated-choice surveys? Health Econ. 2010;19:111-8.
47. Murphy TV, Gargiullo PM, Massoudi MS, et al. Intussusception among infants given an oral rotavirus vaccine. N Eng J Med. 2001;344:564-72.

ADDITIONAL FILE 1: DESCRIPTION OF BOTH STUDIES

Rotavirus DCE

Between January and March 2013, Veldwijk et al. [1] conducted a DCE on parental preferences for rotavirus vaccination of new-borns in the Netherlands. A random sample of 2500 parents with new-borns was selected from the Praeventis database (a national vaccination register in which the vaccination status of all Dutch new-borns is registered) to participate in this study. The DCE questionnaire consisted of nine choice tasks. In the choice tasks, participants were asked to choose between two different rotavirus vaccine scenarios to protect their child from an infection. Each choice task was constructed based on five attributes with either two or three levels (Table 1a). In total, 959 participants completed the questionnaire, of which 202 gave permission to be re-contacted for further research.

Prostate cancer-screening DCE

To investigate men's preferences and trade-offs for prostate cancer screening, de Bekker-Grob et al. [2] conducted a DCE between January and May 2011 among a population-based random sample of 1000 men aged 55 to 75, living in the Rijnmond region of the Netherlands. The DCE questionnaire comprised 16 choice tasks, in which participants were asked to choose between a no screening scenario (opt-out) and two prostate cancer-screening scenarios. Each scenario consisted of five attributes, with each four levels (Table 1b). In total, 459 men responded to the questionnaire and 373 gave permission to be contacted again for additional questions.

REFERENCES

1. Veldwijk J, Lambooij MS, Bruijning-Verhagen P, et al. Parental preferences for rotavirus vaccination in young children: a Discrete Choice Experiment. Vaccine. 2014;32:6277-83.
2. De Bekker-Grob EW, Rose JM, Donkers B, et al. Men's preferences for prostate cancer screening: a discrete choice experiment. Br J Cancer. 2013;108:533-41.

ADDITIONAL FILE 2: EXAMPLE OF CHOICE TASKS

Rotavirus DCE [1]

Imagine that a vaccine against rotavirus infections would become available within the Netherlands. In what situation would you prefer to vaccinate your new-born, situation 1 or situation 2?

	Situation 1	Situation 2
Effectiveness	75%	95%
Frequency of severe side effects	1 in 100,000	1 in 10,000
Protection duration	3 years	6 years
Location	Your General Practitioner	Child Welfare Center
Costs	30 Euro	0 Euro
I would choose for:	☐	☐

Prostate cancer-screening DCE [2]

Thirty-five out of every 1000 deaths among men are caused by prostate cancer. Which alternative do you prefer to reduce your risk of dying from prostate cancer: no screening, screening program 1, or screening program 2? (please, tick one box)

	No screening	**Program 1**	**Program 2**
Amount of men per 1.000 men who will **die** from prostate cancer	**35 deaths**  (0 deaths prevented)	**25 deaths**  (10 deaths prevented)	**18 deaths**  (17 deaths prevented)
Frequency of a blood **test**	**No** blood test	Every **4 years** a blood test	Every **3 years** a blood test
Amount of men per 1,000 men with an increased PSA who receive an **unnecessary biopsy** (= no cancer detected, although the blood test suggested that a blood test was necessary)	**Not applicable**	**400 unnecessary biopsies**  (600 correct biopsies)	**800 unnecessary biopsies**  (200 correct biopsies)
Amount of men per 1,000 treated men who receive an **unnecessary treatment** (= no increase in life expectancy, but there is a risk of urine incontinence and erection problems due to treatment)	**Not applicable**	**0 unnecessary treatments**  (1,000 correct treatments)	**500 unnecessary treatments**  (500 correct treatments)
Out of pocket cost per year during the period of the screening program	**0 euro per year**	**100 euro per year** 	**50 euro per year** 
Which alternative would you choose?	☐	☐	☐

REFERENCES

1. Veldwijk J, Lambooij MS, Bruijning-Verhagen P, et al. Parental preferences for rotavirus vaccination in young children: a Discrete Choice Experiment. Vaccine. 2014;32:6277-83.
2. De Bekker-Grob EW, Rose JM, Donkers B, et al. Men's preferences for prostate cancer screening: a discrete choice experiment. Br J Cancer. 2013;108:533-41.

ADDITIONAL FILE 3: DESCRIPTION OF THE HEALTH LITERACY MEASURES

The prostate cancer-screening cohort was asked to complete the three validated Dutch questions of the Set of Brief Screening Questions (SBSQ-D) of Chew [1] to measure their subjective health literacy. This instrument was already included in the initial rotavirus DCE, these questions were therefore not repeated in the current study. The SBSQ-D contains questions on how often participants need help to read letters from their GP/specialist, how sure participants are that they complete medical forms correctly and how often participants find it difficult to find information about their health. Participants scored these questions on a 5-point Likert scale, from zero to four. An average score of ≤2 indicates inadequate health literacy, while an average score >2 indicates adequate health literacy [1]. The Dutch version of the Newest Vital Sign (NVS-D) was included as an objective measure of health literacy [1] for both groups. To measure the participants' health literacy status, they were asked six questions about an ice cream nutrition label. Participants scored one point for each correctly answered question, with a maximum of 6 points. A score of 4-6 indicates adequate health literacy [1].

REFERENCE

1. Fransen MP, Van Schaik TM, Twickler TB, et al. Applicability of internationally available health literacy measures in the Netherlands. J Health Comm. 2011;16Suppl3:134-49.

11

General discussion

The objective of this dissertation was to contribute to the growing field of health-related stated-preference research by addressing research questions that relate to the past (research question 1), the present (research questions 2a-2c) and the future (research questions 3a-3c) of the discrete choice experiment (DCE) methodology. In this final chapter, the main findings are described first. The methodological considerations are described next, followed by the interpretation of the main findings and directions for further research. This chapter concludes with recommendations for health policy and recommendations for DCE researchers.

MAIN FINDINGS

Part I: The past - review of the literature

1. What are recent practices and trends, including progress in methodology, in applications of health-related DCEs?

The systematic literature review of health-related DCE studies published between 2009 and 2012 (**Chapter 2**) shows that the use of DCEs continues to grow, as does the scope of applications across an expanding range of countries. We found that qualitative methods are less often used to inform attribute selection, compared to the period 2001-2008. The trend towards the use of D-efficient designs and the use of more flexible econometric methods, including mixed logit and latent class models, continued. There has been a shift from paper DCE surveys towards computer-administered surveys and the use of Internet panels.

Part II: The present - three state of the art applications

2a. What are preferences of European citizens for vaccination programmes during future pandemics?

The qualitative exploration of public opinion and attitudes in three European countries (The Netherlands, Sweden and Poland; **Chapter 3**) showed that participants would base their vaccination decision on trade-offs between perceived benefits and barriers of the vaccine also taking into account the seriousness of the new outbreak. Except for those having chronic diseases, participants expected a low infection risk. Participants displayed concerns about vaccine safety due to the limited available time to produce and test vaccines in the acute situation of a new pandemic. Important differences between respondents from different countries stood out, such as previous vaccination experiences, the degree of adherence to social norms, and the degree of trust in health authorities. The findings of this study served as input for the attribute and level selection of the DCE on preferences for pandemic

vaccinations. **Chapter 4** describes this DCE application on pandemic vaccinations that was conducted in four European countries (The Netherlands, Spain, Sweden and Poland). In the case of a severe pandemic scenario, vaccine effectiveness was the most important characteristic determining vaccination preferences in all four countries, followed by the body that advises the vaccine. In Sweden, the advice of family and/or friends and the advice of physicians strongly affect vaccine preferences, in contrast to Poland and Spain, where the advice of (international) health authorities was more decisive. Irrespective of pandemic scenario or vaccination programme characteristics, the predicted vaccination uptakes were lowest in Sweden, and highest in Poland.

The DCE data of the Netherlands was further used to study within country differences in preferences for pandemic vaccinations (**Chapter 5**). Substantial preference heterogeneity was found. Females and individuals who stated that they were never in favour of vaccination made different trade-offs than males and individuals who stated that they were (possibly) willing to get vaccinated.

2b. What trade-offs do consumers make between basic health plan characteristics in the Dutch health insurance market?

Chapter 6 describes the DCE application on basic health plans. Being able to choose a care provider freely (and associated reimbursement level) was by far the most decisive characteristic for respondents aged over 45, those with chronic conditions, and those with a gross income over €3000/month. Although choice of care provider (and associated reimbursement level) was also an important characteristic for those younger, healthier, and with a lower income, monthly premium was their most important determinant of choice. Health insurers with a focus on contracting high quality providers were preferred over health insurers with a focus on lower-priced providers, although WTP estimates were generally low. Half of the sample would never choose health plans with restricted provider choice. However, a premium discount up to €15/month by restricted health plans might motivate especially younger, healthier, and less wealthy consumers to choose these plans.

2c. What are personal health records preferences of potential users?

Chapter 7 describes the DCE application on personal health records (PHRs). Three latent classes with different preference structures were identified: “refusers”, “eager adopters” and “reluctant adopters”. Those with one or more chronic diseases were significantly more likely to belong to the eager adopter’s class. The uptake of the reluctant adopters varied between 4% and 68% depending on PHR characteristics as varied in our DCE, while the uptake of the refusers was always below 9% and the uptake of the eager adopters was always above 91%. The facilitating party for storage of PHR was the most decisive aspect for

the eager and the reluctant adopters, while the costs were most decisive for the refusers. Across all classes, care providers and independent organisations were the most preferred facilitators of data storage.

Part III: The future - methodological studies

3a. Does the inclusion of an "opt-out" instead of a "neither" alternative affect DCE results? Does the inclusion of a "status quo" alternative in addition to an "opt-out" alternative affect DCE results in markets where there is a status quo?

The methodological study described in **Chapter 8** shows that the use of different unforced choice formats affects marginal utilities and welfare estimates and hence the conclusions that will be drawn from the DCE. Only a minority of respondents stated that they interpreted the neither alternative as analysts normally code it (zero option).

3b. Does the choice of DCE administration mode (paper versus online) affect the result of a DCE?

The methodological study described in **Chapter 9** shows that the median response time was shorter for the online sample compared to the paper sample, and that a smaller proportion of respondents from the online sample was satisfied with the number of choice sets. In addition, a larger proportion of respondents from the online sample, compared to the paper sample, failed the monotonicity axiom. We found some significant between-sample differences in choice probabilities for the last five choice sets. Although some WTP estimates were higher for the online sample, the elicited preferences for basic health plan characteristics were similar between both modes of administration. We found no indication that online surveys yield inferior results compared to paper based surveys, while the price per respondent is generally lower for online surveys.

3c. How do respondents complete the choice sets in a DCE?

The interview study described in **Chapter 10** revealed that the majority of the participants preferred better over worse risk attribute levels and answered the control questions regarding their understanding of the numerical values of the risk attributes correctly. At the same time, a large proportion of the participants was not able to repeat the exact definition of the risk attributes as explained to them in the introduction of the survey. The majority of the participants based their decision on three or more attributes by trading them against each other, which implies complex decision strategies and is in accordance with the continuity axiom. However, lower educated and less literate participants tended to mention less than three attributes when motivating their decision and used trading off between attributes less often as a decision-making strategy.

METHODOLOGICAL CONSIDERATIONS

Similar to all research methods, DCEs have specific benefits, and some restrictions. In the following sections, methodological considerations of the studies in this dissertation will be discussed.

Part I: The past - review of the literature
In the interests of ensuring comparability with data from the earlier review that included studies published between 2001-2008 [1], the searches to identify the health-related DCEs published between 2009 and 2012 were restricted to PubMed only (**Chapter 2**).

Part II: The present - three state of the art applications
Conclusions drawn from the focus group study (**Chapter 3**) should be considered with some caution as the findings are based on a small number of individuals, and may therefore not be generalizable to populations at large. However, this qualitative research served as important input for the selection of attributes and levels of the DCE. In addition, it contributed to the interpretation of the DCE results.

Four important issues need to be considered when interpreting the findings from the three DCE applications included in this dissertation:

First, because a discrete choice model depends completely on the attributes that are included in the DCE, exclusion of important attributes might lead to biased part-worth utilities and inaccurate welfare measures [2-4]. For all the applications in this dissertation, we aimed to select the most important attributes and corresponding levels by carefully studying the literature, interviewing experts, conducting focus group discussions with the target population of the DCE and pilot testing of the DCE. However, we cannot exclude the option that we missed an attribute or a level that would have affected our findings.

Second, in all three applications, we used Internet panels to collect the DCE data. It might be that those panel members that were not interested in the topic did not start the survey, which in turn might have affected estimated preferences. Responder bias can thus not be excluded. We used commercial panels in **Chapter 4, 5, 6 and 9**, and participation rates were comparable to the average rate of the Internet panels we used. We do expect our results to be generalizable since age, gender, level of education and region of our sample are comparable to that of the general population of the Netherlands. The NPCF Internet panel was used to collect the DCE data of **Chapter 7 and 8.** The study sample may not be fully representative for the Dutch general adult population. Respondents were older, more highly educated, and more often had chronic diseases.

Third, another issue is the complexity of the discrete choice task. Due to both the number and the type of attributes and levels that respondents needed to take into account when completing the choice sets (especially due to inclusion of probabilities as attributes in **Chapter 4, 5, 6, 9, and 10**) it can be expected that respondents might have experienced difficulties completing the DCE, which might have influenced the results. In **Chapter 4 and 5**, we were advised by experts in the field of risk communication on how to present the choice sets, we included graphs to demonstrate percentages and rates, and used realistic presentation of attributes (i.e. integers when discussing rates that included humans). Piloting and think-a-loud interviews in the preparation phase, as well as questions that assessed the experienced difficulty of the survey showed that the majority of respondents had no problems with completing the choice sets. Additionally, the signs of the coefficients were generally consistent with our a priori hypothesis, e.g. a higher vaccine effectiveness would have a positive effect on vaccination, which confirms the theoretical validity of the DCE [3].

Fourth, preferences in DCEs are stated and based on hypothetical scenarios. An advantage of stated over revealed preference data is that the uptake of new products with new attributes can be modelled within a controlled experiment [5], see for example **Chapter 4** in which we were able to estimate preferences for and uptake of vaccinations in future pandemics. The DCE-based predicted uptake could be used in economic evaluations [6]. However, respondents face no direct consequences of their choices in a DCE [7] and might have given socially desirable responses. DCEs could therefore suffer from hypothetical bias, for example, the WTP estimates and/or predicted uptake rates might differ from real life situations. The external validity of the DCE method has been studied in some other health related contexts, and results are encouraging with respect to prediction of preferences on an aggregate level [8-10]. However, the external validity of DCEs is an under-researched topic, and one of the reasons that has been stated for this is the lack of available revealed-preference data [11]. The DCE on health plan preferences described in **Chapter 6** might be a suitable application to test the external validity of DCEs, given that respondent's choices for a certain health plan can easily be reported by respondents.

Part III: The future - methodological studies

Because of data-collection software limitations, we were not able to present respondent specific status quo (SQ) levels in the choice sets and to pivot the attribute levels of the hypothetical alternatives around the SQ (e.g. xx % more expensive than your SQ). Our finding that SQ should be included in addition to opt-out (**Chapter 8**) might have been even more pronounced if respondents could have seen their individual SQ levels in the choice sets.

Although we successfully controlled for well-known confounders that may distort the comparison of paper survey and online responses [12, 13] in **Chapter 9,** there might still be differences in unobserved respondent characteristics that could have biased our results.

The sample frame we used (Internet), might have affected our results because Internet panel members are experienced to completing surveys online, not on paper. In addition, these Internet panel members have actively signed-up to complete surveys. Results might have shown more differences between administration modes if we would have used other samples of respondents, e.g. random digit dialling versus an Internet panel. The generalizability of our findings to other respondent groups (e.g. elderly) or other topics may be limited. More research on the generalizability of the findings to other topics and target groups is thus needed.

The two DCEs that we used as case studies for the study in **Chapter 10** were quite complex, because each included two risk attributes. It is commonly known that participants perceive the interpretation of such risk attributes as more difficult than qualitative attributes [14, 15]. Difficulties in interpreting attribute levels and making decisions might therefore be more pronounced in this study compared to DCEs that include no or less risk-related attributes. However, since most health-related decisions include risk information, the case studies used for this study may be representative for many DCEs within a healthcare context. Although efforts were made to mimic non-lab choice situations during these interviews, the fact that the interviewers were present during DCE completion might have influenced how participants completed the choice sets. Participants therefore might have been more committed to completing the DCE. As a result, we might have overestimated the number of participants that acts in accordance with the tested assumptions. The sample size of 70 is relatively large for an interview study, at the same time; this sample size is too small to draw any conclusions based on statistical testing. However, the trends in the findings and the agreement of the current findings with the existing literature related to educational level and health literacy (non-DCE studies) provide face validity for the current study results.

INTERPRETATION OF EMPIRICAL FINDINGS AND DIRECTIONS FOR FURTHER RESEARCH (PART II - THE PRESENT)

Preferences for pandemic vaccinations

The studies described in **Chapter 3, 4 and 5** show that seriousness of a pandemic influences vaccination uptake dramatically. In order to increase pandemic vaccination coverage, it is essential that susceptible people feel susceptible and perceive the pandemic as a serious threat. This can be achieved, for example, by avoiding conflicting messages and information overload [16-18]. To increase attention to information that is being communicated in pandemics it is important to acknowledge any uncertainty about this pandemic information, and to highlight that this uncertainty is common during almost all pandemics. Visualize risk using a heat map when communicating the number of cases and deaths in each country, and use a label that is connected to the animal vector (e.g. Swine flu) or that sounds unusual, to increase interest in more information [19].

The vaccination program attributes that can be influenced by policy makers directly are out-of-pocket costs and how/what to communicate. As our results show that by whom a vaccine is advised had a different effect on uptake in the included countries, it is important that during future pandemics the responsible authorities align with other important stakeholders in the country and communicate in a coordinated manner. In addition, it is important to build trust in the pre-outbreak phase, maintain trust during outbreaks and, if necessary, restore or further develop trust after the pandemic ends [20, 21].

Swedish participants indicated that their experiences during the Influenza A/H1N1 2009 pandemic would reduce their tendency to accept future vaccination advice **(Chapter 3)**. These discussions may be rooted in the Swedish government having signed a contract with a pharmaceutical company to buy pandemic Influenza A/H1N1 vaccines years before the outbreak [22] and the high incidence rates of narcolepsy following the Influenza A/H1N1 2009 pandemic, suggesting an association with vaccination [23, 24]. The seasonal Influenza vaccination coverage in Sweden decreased since the Influenza A/H1N1 2009 pandemic; it was 65.8% in 2008-2009 but decreased to 44.3% in 2012-2013 [25]. Combined with the Polish participants being proud that their Minister of Health had not bought vaccines during the Influenza A/H1N1 2009 pandemic, these findings confirm what Börjesson et al concluded in 2013 [26]; previous experiences with outbreak situations play a crucial role in public opinions and future behaviour. The focus group study highlights that outbreak experiences differ between countries in many dimensions: with regard to cultural differences, with regard to government policies, and with regard to vaccination side effects (narcolepsy in Sweden). These differences stress the need to adapt communication strategies to local circumstances.

Given potential differences in cultures, government policies, trust in health authorities, and/or previous experiences with outbreaks, it may be expected that preferences also differ between countries within the same European region. It might therefore be useful to conduct the same DCE in other European countries within the same European region as well. Future research could use the available DCE data to study differences in preferences within Spain, Sweden and Poland as well. Additionally, future research could focus on subgroups of the population, such as healthcare workers or under-vaccinated groups.

Our studies show that the availability of an effective pandemic vaccine is of paramount importance in order to reach certain coverage levels. Unfortunately, such a highly effective vaccine might not be available due to the crisis-situation that is inherent to a pandemic, or proof that the vaccine is effective might be lacking, as time is usually limited and because it is not known in advance which virus will cause a next pandemic [27]. In addition, due to contracts or limited availability of vaccines, there are usually only one or two different vaccines available for policymakers to choose from. Therefore, other preventive measures such as quarantine, and antiviral drugs might be helpful to limit the spread of the virus during

the first phase of an influenza pandemic [28]. Further research into preferences for other preventive measures, and differences herein across European countries is recommended.

Preferences for basic health insurance

A possible explanation for the dominance of the choice of care providers (and associated reimbursement level) attribute that we found in **Chapter 6** might be that respondents prefer to choose healthcare providers regardless of financial considerations. In addition, the dominance might partly be related to an objection to the government proposal allowing insurers to set a lower level of reimbursement of care by non-contracted providers. At the time of data-collection, this proposal was heavily disputed by the medical profession and widely covered by the media.

Our results indicate that consumers do not value the focus of insurers on quality of care highly. This might indicate that consumers do not expect insurers to play an important role as purchasing agents of high quality care. Instead, consumers seem to rely on their general practitioner to select which are the best care providers to go to [29]. If insurers are able to provide objective quality information of the contracted providers that is easy to understand, consumers might make more well considered insurance choices that do include quality of contracted care. As a result, consumers might be more willing to accept less freedom of provider choice. Recently, the Dutch Minister of Health advised the insurers to be more transparent about their contracting practices [30].

Given that characteristics of supplementary health insurance affect the choice for basic health insurance of the majority of the respondents much or very much (**Chapter 6**), it is important that future research explores the trade-offs that respondents make between basic and supplementary health insurance. We showed in **Chapter 8** that if a DCE concerns a market where a SQ exists, the choice sets of a DCE should include a SQ alternative in addition to opt-out. As the health insurance market is such a market where respondents have a SQ, i.e. their current health plan, future DCEs on health plan preferences might want to consider to include a SQ alternative, especially if the aim is to predict uptake rates.

Preferences for personal health records

The study described in **Chapter 7** showed that it is of importance that information campaigns of policy makers as well as of PHR producers are targeted to chronically ill people given our finding that those respondents had a higher probability to belong to the eager adopters. This large class of respondents showed to have great interest in PHRs, irrespective of its characteristics. Given that our sample consisted of people that currently do not have a PHR, they might be willing to have a PHR, but they might not know how to get one, or the PHRs that are currently on the market are not interesting to them. The reasons for this might be an area for further research. Further research may also want to focus on preferences of current PHR users and how their preferences differ from potential PHR users.

It appears to be extremely difficult to increase the uptake of PHRs by creating a better product given our finding that the PHR uptake of only 20% of respondents was influenced by the characteristics of a PHR. The ideal PHR of this group would be a PHR which is facilitated by the care provider and tethered with the GP system. The data would be used for other purposes after permission only and adding own data would be possible at zero costs to the user. Policymakers who aim to expand the use of PHRs will be most successful when healthcare providers and health facilities, or independent organizations facilitate the storage of PHR data, while refraining to include market parties in storage of the data. Low costs, some form of connection with other systems, and the option to upload own data are valued positively by potential users, but these aspects will affect the uptake to a smaller extent. Producers of PHRs need to convince the potential users that they secure the privacy of the PHR data.

INTERPRETATION OF THEORETICAL FINDINGS (PART I - THE PAST & PART III - THE FUTURE)

The decline in the use of qualitative methods to inform attribute selection (**Chapter 2**) is of concern. If the selection of attributes is not properly grounded in qualitative research, important attributes might not be identified, resulting in a less usable and less valid design [4]. If qualitative methods have been used they are usually reported without much detail [31], and it is often not clear how exactly attributes and attributes levels were selected [32]. The process of how certain attributes and levels are selected based on the qualitative research performed is often a black box that involves numerous decisions of the researchers. Coast et al. called for “rigorous and transparent” reporting of the attribute selection process [32]. This might even result in writing a separate paper on the process of attribute and level development. A good example of such practice is a paper by Abiiro et al. in which the systematic process of developing attributes and attribute levels for a DCE on micro health insurance is reported [33].

Based on the methodological study presented in **Chapter 8**, health economists engaged in preference elicitation tasks are urged to carefully consider how to design the unforced choice experiment. We recommend to use opt-out instead of neither, as neither represents an unspecified alternative that is interpreted differently by respondents, and seldom as a true opt-out alternative. We further recommend to include both an opt-out and a SQ alternative in markets with a SQ if this best mimics the real market situation. Given that DCE designs are increasingly web-based and interactive, initial SQ questions can easily feed into tailored DCE questions presented to the individual respondent. The standard argument for not including respondents’ own SQ levels is that such products are often marketed as non-differentiated products, and that consumers’ current products are therefore hard to

identify. Further, collecting information on the SQ alternative requires that consumers can correctly identify which attribute levels describe their usual product [34].

DCE data is increasingly being collected online (**Chapter 2**). We find no indication that online surveys yield inferior results compared to paper based surveys, while the price per respondent is generally lower for online surveys (**Chapter 9**). Savage et al. [35] compared respondents' learning and fatigue between mail and online DCE administration. They found that online respondents who completed the DCE suffered from fatigue or boredom, while mail respondents did not and where therefore more consistent in their answers [35]. Hess et al. found more evidence of learning (in terms of utility scale) than fatigue in a number of DCE applications, while no relation between results and mode of administration was found (computer aided personal interviews versus online) [36]. We found some significant between-sample differences in choice probabilities for the last five choice sets, which supports the recommendation of Savage et al. that researchers might need to include fewer choice sets per respondents when collecting DCE data online, compared to when collecting DCE data using paper surveys. In order to collect the same number of choice observations, researchers might want to increase the number of respondents that participate in the online DCE. This may be feasible since the costs for online data-collection are usually lower. The majority of the DCEs published between 2009-2012 included 9 to 16 choice sets (**Chapter 2**). However, there is no consensus on the 'right' number of choice sets in the DCE literature. Watson et al. found mixed effects of the number of choice sets on response rates in a meta-regression analysis on postal DCE surveys: response rates are higher when a DCE includes 3-7 or more than 8 choice sets, compared to 8 choice sets [37].

Results of the interview study described in **Chapter 10** indicate that some respondents have difficulties understanding all the information that is provided to them in the DCE, that they do not always use complex decision strategies to make their choice and therefore do not always act in accordance to the continuity axiom. This was most prominent in respondents with a lower educational level, higher age and lower health literacy status. Additional actions should therefore be undertaken to ensure that participants understand the choice task and complete the DCE as presumed. For example, DCE surveys among older and/or less literate populations could be conducted in mini-labs, where participants complete the DCE in the presence of a researcher. Recent research shows that allowing for attribute non-attendance (ANA, ignoring one or more attributes) in the choice models, for example by using latent class models, positively affects the model fit [38, 39]. Another study recommended to also take attribute-level non-attendance into account when modelling choice data because respondents process each level differently [40]. Attribute level overlap was identified as an effective strategy to improve attribute attendance and reduce task complexity in DCEs [41].

In order to exploit the full potential that DCEs yield, it is essential that policy makers pick up the generated evidence. However, the use of available evidence on patient preferences in, for example clinical practice guidelines development and coverage decisions [42], is currently limited. It might therefore be worthwhile to increase interactions between researchers and policy makers [43], by already involving policy makers - or other end users of the DCE results – when designing the DCE. In addition, researchers need to actively approach policy makers to present their research findings and to discuss the implications for health policy. To increase societal impact, researchers might also want to consider publishing the DCE results in journals that are easily accessible for policy makers, in addition to scientific journals.

DIRECTIONS FOR FURTHER METHODOLOGICAL DCE RESEARCH (PART I - THE PAST & PART III - THE FUTURE)

The review of published health-related DCE studies (**Chapter 2**) needs to be updated to the most recent years to continuously monitor current practices and to identify areas for methodological research. When doing so, searches should not be restricted to PubMed only. In addition, it is important to follow progress in methodology in other fields that have a longer tradition of using DCEs than health economics, for example transport and environmental economics [44], closely and to adapt and apply their best practices to the health economics field.

The study in **Chapter 8** is the first in health that examines the use of different choice formats in DCEs. In addition, it is the first study in the field of discrete choice modelling that compared opt-out with neither, and explored the effects of including a SQ in addition to an opt-out. More research is clearly needed to test the replicability of the findings. In addition, future research could investigate if respondents indeed perceive the opt-out alternative as the option not to buy the good or not to choose the service.

Previous research shows no mode effects between a mobile app and computer web survey administration [45]. Although we did not receive any negative feedback hereof in the study presented in **Chapter 9**, it can be questioned whether the choice sets of a DCE are fully visible on the relatively small smartphone screens. Future studies might therefore want to focus on the effects of completing a DCE on different digital devices, such as laptops, tablets, and smartphones. Due to its technical possibilities, electronic data-collection offers additional benefits over paper data-collection. For example, explanations using video might improve the survey understanding of lower literate respondents [46]. It would be worthwhile if future DCE studies test whether such features aid in the understanding of the DCE. Electronic data-collection also provides more control and flexibility over the information that is presented in a survey [47] compared to paper data-collection and can

therefore more easily facilitate methodologic DCE studies that incorporate split samples with different presentation formats, e.g. when studying ordering effects [48]. An additional benefit of collecting data online is that response times can be logged. Earlier work shows that the time respondents take when answering the choice sets can be seen as a proxy for respondent engagement, and needs to be taken into account when modelling choice data, as it influences the mean and the variances of random parameter distributions [49].

We used the think aloud method to study how respondents complete a DCE in **Chapter 10**. Future research could use eye-tracking to explore visual attention sequences and underlying decision processes, as well as reading strategies regarding for instance the opt-out option. Additionally, eye-tracking research will also provide insights in the potential discrepancy between the way respondents complete a DCE with or without thinking aloud. Eye-tracking seems to be a promising method to investigate respondents' behaviour when completing a DCE, see for example [50, 51].

RECOMMENDATIONS FOR HEALTH POLICY

1. During future pandemics, responsible authorities should align with important stakeholders in a country and communicate in a coordinated manner.
2. Adapt pandemic preparedness plans and communication strategies to local circumstances.
3. A premium discount of up to €15/month on restricted health plans might motivate especially younger, healthier and less wealthy consumers to choose these plans.
4. Policymakers who aim to expand the use of personal health records (PHRs) in the Netherlands will be most successful when healthcare providers or independent organizations facilitate the storage of PHR data, while refraining from including market parties in the storage of such data.
5. Target PHR information campaigns to chronically ill people.

RECOMMENDATIONS FOR DCE RESEARCHERS

1. Use qualitative methods (such as focus group discussions) to inform attribute and level selection, and report more details on this selection process.
2. Follow progress in methodology in research fields that have a longer tradition of using the DCE methodology.
3. Involve policy makers, or other end users of the DCE results, when designing the DCE. In addition, approach policy makers actively to present research findings and to discuss the implications for health policy. Consider publishing the DCE results in journals that are easily accessible for policy makers.
4. Use “opt-out” instead of “neither” in unforced DCEs.
5. Include a “status quo” alternative in addition to “opt-out” in markets where a status quo exists.
6. We find no indication that online surveys yield inferior results compared to paper based surveys. However, especially in online DCEs, researchers might want to include fewer choice sets per respondent to avoid fatigue or boredom of respondents.
7. Include age, education and health literacy measures in the DCE survey as well as in the analysis of DCE data. Until options to correct DCE responses for possible differences in demographic characteristics become common practice, researchers should at least describe these measures in their population and explain their possible effects on the results.

REFERENCES

1. De Bekker-Grob EW, Ryan M, Gerard K. Discrete choice experiments in health economics: a review of the literature. Health Econ. 2012;21:145-72.
2. Kjaer T. A review of the discrete choice experiment-with emphasis on its application in health care. Syddansk Universitet Odense, Denmark; 2005.
3. Lancsar E, Louviere J. Conducting discrete choice experiments to inform healthcare decision making: a user's guide. Pharmacoeconomics. 2008;26:661-77.
4. Kløjgaard ME, Bech M, Søgaard R. Designing a stated choice experiment: the value of a qualitative process. JOCM. 2012;5:1-18.
5. Mark TL, Swait J. Using stated preference and revealed preference modeling to evaluate prescribing decisions. Health Econ. 2004;13:563-73.
6. Terris-Prestholt F, Quaife M, Vickerman P. Parameterising User Uptake in Economic Evaluations: The role of discrete choice experiments. Health Econ. 2016;25 Suppl 1:116-23.
7. Harrison GW. Making choice studies incentive compatible. In: Kanninen BJ, editor. Valuing environmental amenities using stated choice studies. Springer; 2006. p. 67-110.
8. Kesternich I, Heiss F, McFadden D, Winter J. Suit the action to the word, the word to the action: Hypothetical choices and real decisions in Medicare Part D. J Health Econ. 2013;32:1313-24.
9. Linley WG, Hughes DA. Decision-makers' preferences for approving new medicines in Wales: a discrete-choice experiment with assessment of external validity. Pharmacoeconomics. 2013;31:345-55.
10. Lambooij MS, Harmsen IA, Veldwijk J, et al. Consistency between stated and revealed preferences: a discrete choice experiment and a behavioural experiment on vaccination behaviour compared. BMC Med Res Methodol. 2015;15:19.
11. Lancsar E, Swait J. Reconceptualising the External Validity of Discrete Choice Experiments. Pharmacoeconomics. 2014;32:951-65.
12. Barentsz MW, Wessels H, van Diest PJ, et al. Tablet, Web-Based, or Paper Questionnaires for Measuring Anxiety in Patients Suspected of Breast Cancer: Patients' Preferences and Quality of Collected Data. J Med Internet Res. 2014;16.
13. Diment K, Garrett-Jones S. How demographic characteristics affect mode preference in a postal/web mixed-mode survey of Australian researchers. Soc Sci Comp Rev. 2007;25:410-7.
14. Galesic M, Garcia-Retamero R. Statistical numeracy for health: a cross-cultural comparison with probabilistic national samples. Arch Intern Med. 2010;170:462-8.
15. Waters EA, Weinstein ND, Colditz GA, et al. Formats for improving risk communication in medical tradeoff decisions. J Health Comm. 2006;11:167-82.
16. Larson HJ, Cooper LZ, Eskola J, et al. Addressing the vaccine confidence gap. Lancet. 2011;378:526-35.
17. Lin L, Savoia E, Agboola F, et al. What have we learned about communication inequalities during the H1N1 pandemic: a systematic review of the literature. BMC Public Health. 2014;14:484.
18. Van der Linden SL, Clarke CE, Maibach EW. Highlighting consensus among medical scientists increases public support for vaccines: evidence from a randomized experiment. BMC Public Health. 2015;15:1207.
19. Effective Communication in Outbreak Management (ECOM): recommendations for communicating about influenza pandemics. Available from: https://sites.google.com/site/ecomwp8/home. [Last accessed 29 04 2016].
20. Paek HJ, Hilyard K, Freimuth VS, et al. Public support for government actions during a flu pandemic: lessons learned from a statewide survey. Health Promot Pract. 2008;9:60S-72S.
21. Gesser-Edelsburg A, Mordini E, James JJ, et al. Risk Communication Recommendations and Implementation During Emerging Infectious Diseases: A Case Study of the 2009 H1N1 Influenza Pandemic. Disaster Med Public Health Prep. 2014:1-12.
22. Widgren K, Magnusson M, Hagstam P, et al. Prevailing effectiveness of the 2009 influenza A(H1N1) pdm09 vaccine during the 2010/11 season in Sweden. Euro Surveill. 2013;18:20447.

23. Barker CI, Snape MD. Pandemic influenza A H1N1 vaccines and narcolepsy: vaccine safety surveillance in action. Lancet Infect Dis. 2014;14:227-38.
24. Wijnans L, Lecomte C, de Vries C, et al. The incidence of narcolepsy in Europe: before, during, and after the influenza A(H1N1)pdm09 pandemic and vaccination campaigns. Vaccine. 2013;31:1246-54.
25. Seasonal influenza vaccination programme country profile: Sweden (2012-13 Season). European Centre for Disease Prevention and Control (ECDC), 2014. Available from: http://ecdc.europa.eu/en/publications/Report%20Assets/seasonal-vaccination-coverage-in-europe-2012-13/Seasonal-Influenza-Vaccination-Programme-Country-Profile-Sweden.pdf.
26. Börjesson M, Enander A. Perceptions and sociodemographic factors influencing vaccination uptake and precautionary behaviours in response to the A/H1N1 influenza in Sweden. Scand J Public Health. 2014;42:215-22.
27. Evolution of a pandemic: A(H1N1) 2009, April 2009 - August 2010. World Health Organization (WHO), 2013. Available from: http://apps.who.int/iris/bitstream/10665/78414/1/9789241503051_eng.pdf?ua=1.
28. Longini IM, Jr., Nizam A, Xu S, et al. Containing pandemic influenza at the source. Science. 2005;309:1083-7.
29. Rademakers J, Nijman J, Brabers AE, et al. The relative effect of health literacy and patient activation on provider choice in the Netherlands. Health Policy. 2014;114:200-6.
30. [Letter from the Minister to the House of Representatives 'Transparancy health insurers']. Kamerbrief 'Transparantie zorgverzekeraars' van Minister Schippers aan de Tweede Kamer. 2015. Available from: http://www.rijksoverheid.nl/documenten-en-publicaties/kamerstukken/2015/06/30/kamerbrief-over-transparantie-zorgverzekeraars.html. [Last accessed 24 04 2016]. *In Dutch.*
31. Coast J, Horrocks S. Developing attributes and levels for discrete choice experiments using qualitative methods. J Health Serv Res Policy. 2007;12:25-30.
32. Coast J, Al-Janabi H, Sutton EJ, et al. Using qualitative methods for attribute development for discrete choice experiments: issues and recommendations. Health Econ. 2012;21:730-41.
33. Abiiro GA, Leppert G, Mbera GB, et al. Developing attributes and attribute-levels for a discrete choice experiment on micro health insurance in rural Malawi. BMC Health Serv Res. 2014;14:235.
34. Cameron TA, DeShazo J, Johnson EH. Scenario adjustment in stated preference research. JOCM. 2011;4:9-43.
35. Savage SJ, Waldman DM. Learning and fatigue during choice experiments: a comparison of online and mail survey modes. J Appl Econom. 2008;23:351-71.
36. Hess S, Hensher DA, Daly A. Not bored yet–revisiting respondent fatigue in stated choice experiments. Transp Res Part A Policy Pract. 2012;46:626-44.
37. Watson V, Becker F, de Bekker-Grob EW. Discrete Choice Experiment Response Rates: A Meta-analysis. Health Econ. 2016; [Epub ahead of print].
38. Lagarde M. Investigating attribute non-attendance and its consequences in choice experiments with latent class models. Health Econ. 2013;22:554-67.
39. Hole AR, Norman R, Viney R. Response Patterns in Health State Valuation Using Endogenous Attribute Attendance and Latent Class Analysis. Health Econ. 2016;25:212-24.
40. Erdem S, Campbell D, Hole AR. Accounting for Attribute-Level Non-Attendance in a Health Choice Experiment: Does it Matter? Health Econ. 2015;24:773-89.
41. Jonker MF, Donkers B, de Bekker-Grob EW, et al. Using overlap (and color coding) to improve attribute non-attendance and response efficiency in DCEs. 2016; [Submitted].
42. Utens CM, van der Weijden T, Joore MA, Dirksen CD. The use of research evidence on patient preferences in pharmaceutical coverage decisions and clinical practice guideline development: exploratory study into current state of play and potential barriers. BMC Health Serv Res. 2014;14:540.
43. Hanney SR, Gonzalez-Block MA, Buxton MJ, Kogan M. The utilisation of health research in policy-making: concepts, examples and methods of assessment. Health Res Policy Syst. 2003;1:2.
44. Hoyos D. The state of the art of environmental valuation with discrete choice experiments. Ecol Econ. 2010;69:1595-603.

45. Wells T, Bailey JT, Link MW. Comparison of smartphone and online computer survey administration. Soc Sci Comput Rev. 2014;32:238-55.

46. Wright KB. Researching Internet-Based Populations: Advantages and Disadvantages of Online Survey Research, Online Questionnaire Authoring Software Packages, and Web Survey Services. J Comput-Mediat Comm. 2005;10:00.

47. Tingling P, Parent M, Wade M. Extending the capabilities of Internet-based research: lessons from the field. Internet Research. 2003;13:223-35.

48. Kjaer T, Bech M, Gyrd-Hansen D, et al. Ordering effect and price sensitivity in discrete choice experiments: need we worry? Health Econ. 2006;15:1217-28.

49. Rose JM, Black IR. Means matter, but variance matter too: Decomposing response latency influences on variance heterogeneity in stated preference experiments. Market Lett. 2006;17:295-310.

50. Spinks J, Mortimer D. Lost in the crowd? Using eye-tracking to investigate the effect of complexity on attribute non-attendance in discrete choice experiments. BMC Med Inform Decis Mak. 2016;16:14.

51. Balcombe K, Fraser I, McSorley E. Visual attention and attribute attendance in multi-attribute choice experiments. J Applied Econom. 2015;30:447-67.

Summary

Samenvatting

SUMMARY

Patient-centeredness, patient empowerment, shared decision-making and self-management are fundamental elements of current health policy in the Netherlands. The involvement of consumers in health policy decision-making is being encouraged more and more. By involving the public, health policies will better reflect public preferences. Preferences can have an impact on the acceptance of and satisfaction with products, services or interventions, as well as on outcomes.

Although numerous research techniques are available to elicit public preferences for health-related topics, this dissertation focuses specifically on the discrete choice experiment (DCE) methodology to quantify public preferences. A DCE is a survey-based stated-preference elicitation technique: respondents are asked to state their preference for a product, service, or intervention in hypothetical choice situations (called choice sets). A DCE survey typically consists of a series of choice sets, each concerning a discrete choice between two or more alternatives. The word 'discrete' indicates that respondents can only choose one of the alternatives. The presented alternatives are decomposed into characteristics (called attributes). Alternatives are distinguished from one-another by the systematic variation in values of the characteristics (called attribute levels). In the introductory chapter (**Chapter 1**) relevant concepts for DCE are introduced.

DCEs originate from mathematical psychology and are standard practice in the field of marketing, transportation and environmental economics. Since the first DCE in health economics was published in 1990 the body of scientific literature is increasing rapidly. This dissertation aimed to contribute to the growing field of health-related stated-preference research by addressing research questions that relate to the past, the present and the future of the DCE methodology. The research questions as well as the outline of this dissertation are also described in **Chapter 1**.

Part I: The past - review of the literature

Chapter 2 gives an overview of recent health-related DCE practices, including progress in methodology since two previously published reviews (covering the periods 1990-2000 and 2001-2008). We identified 179 health-related DCEs that were published between 2009 and 2012. A review of these studies shows that qualitative methods (such as focus groups) are less often used to inform attribute selection. There has been a growing trend towards the use of D-efficient designs and the use of more flexible econometric methods, including mixed logit and latent class models. Additionally, there has been a shift from paper DCE surveys towards computer-administered surveys and the use of Internet panels.

Part II: The present - three state of the art applications

The second part of this dissertation includes a qualitative study and three DCE applications on: pandemic vaccinations, health insurance and personal health records.

The first three chapters of this part pertain to studies carried out in the context of the project 'Effective Communication in Outbreak Management: development of an evidence-based tool for Europe (ECOM)'. **Chapter 3** describes a qualitative exploration of public opinions and attitudes on pandemics and vaccinations across three European countries (The Netherlands, Sweden and Poland). In each country, two focus group discussions were conducted with 6-9 participants per discussion. We found that participants would base their vaccination decisions on trade-offs between perceived benefits and downsides to getting the vaccine (such as side effects), also taking into account the seriousness of the new outbreak. Except for participants with chronic diseases, participants expected a low infection risk. However, overall participants displayed concerns about vaccine safety due to the limited time available to produce and test vaccines in the acute situation of a new pandemic. Important differences between participants from different countries were seen based on previous vaccination experience, the degree of adherence to social norms and the degree of trust in health authorities. The findings from this study served as input for the attribute and level selection of the first DCE application on preferences for pandemic vaccinations described in **Chapter 4 and 5**. A comparison of pandemic vaccinations preferences between four European countries (The Netherlands, Spain, Sweden and Poland) is presented in **Chapter 4**. Internet panel members, nationally representative in terms of age, sex, educational level and region (N=2,068) completed an online DCE survey in 2013. In the case of a severe pandemic scenario, vaccine effectiveness was the most important characteristic that determined vaccination preferences in all four countries, followed by the body that advised on the vaccine. In Sweden, the advice of family and/or friends and the advice of physicians strongly affected vaccine preferences, in contrast to Poland and Spain, where the advice of (international) health authorities was more decisive. Irrespective of pandemic scenario or vaccination programme characteristics, the predicted vaccination uptakes were lowest in Sweden, and highest in Poland. The DCE data of **Chapter 4** was further used to study differences in preferences for pandemic vaccinations within the Netherlands (**Chapter 5**). Females and individuals who stated that they were never in favour of vaccination made different trade-offs when deciding on whether they would get vaccinated in case of a pandemic than males and individuals who stated that they were (possibly) willing to get vaccinated.

Chapter 6 presents the second DCE application of this part, in which consumer trade-offs in health insurance plan characteristics in the Netherlands were explored. Potential differences in these trade-offs on the basis of age, health status and income were also explored. A representative sample (N=533) of the Dutch adult population completed the online DCE

survey during the 2015 annual period in which persons can select a new health plan. We found that being able to choose a care provider freely (and associated reimbursement level) was by far the most decisive characteristic for health plan choices for respondents aged over 45, those with chronic conditions, and those with a gross income over €3000/ month. Although choice of care provider (and associated reimbursement level) was also an important characteristic for those younger, healthier and with a lower income, monthly premium was the most important determinant in their choice. Health insurers with a focus on contracting high quality providers were preferred over insurers with a focus on lower-priced providers, although willingness-to-pay estimates were generally low. Half of the sample would never choose health plans with restricted provider choice and lower level of reimbursement of healthcare costs. However, a premium discount of up to €15/month on restricted health plans might motivate especially younger, healthier and less wealthy consumers to choose these plans.

Chapter 7 concerns the third DCE application on potential user preferences for personal health records (PHRs). A PHR is an electronic application through which health information can be accessed, managed and shared. Opposed to the clinician's medical record, PHRs are managed by patients. The DCE survey was completed in December 2015 and January 2016 by 1,443 Internet panel members of the Dutch Federation of Patients and Consumer Organizations (NPCF). Three subgroups with different preference structures were identified: "eager adopters", "reluctant adopters" and "refusers". Those with one or more chronic disease were significantly more likely to belong to the eager adopter class. The facilitating party for storage of PHR was the most decisive aspect for the eager and the reluctant adopters, while the costs were most decisive for the refusers. Across all classes, care providers and independent organisations were the most preferred facilitators of data storage. The uptake of the eager adopters was always above 91%, the uptake of the reluctant adopters varied between 4% and 68% depending on PHR characteristics, and the uptake of the refusers was always below 9%.

Part III: The future - methodological studies

The third part of the dissertation describes three methodological DCE studies.

DCEs often include unforced choice situations in which respondents are allowed not to choose any of the hypothetical alternatives through inclusion of an "opt-out" (choose not to buy the good or not to choose the service), "neither" (choose neither of the alternatives) and/or "status quo" alternative (choose your current alternative over the hypothetical alternatives). In the study presented in **Chapter 8,** we aimed to investigate the implications of using different unforced choice formats in different settings. The DCE application on personal health records (**Chapter 7**) served as our case. We show that the choice in these design features of a DCE matters, and should receive ample attention amongst researchers

applying DCEs. We recommend to use opt-out instead of neither, as neither represents an unspecified alternative that is interpreted differently by respondents, and seldom as a true opt-out alternative. We further recommend to include both an opt-out and a status quo alternative in contexts with a status quo if this best mimics the real situation.

It is increasingly common to collect DCE data via computers, and more particularly, online. In the study presented in **Chapter 9** we assessed whether the mode of DCE administration (either via pen-and-paper or online) affects the outcomes of the DCE. The DCE application on basic health insurance served as our case (**Chapter 6**). We found no indication that online surveys yield inferior results compared to paper based surveys, while the price per respondent is generally lower for online surveys.

In order to make valid inferences about individuals' preferences, it is essential that researchers understand how participants complete the choice sets in a DCE. During structured interviews, we explored how 70 participants evaluated and completed four choice sets aloud. Hereafter, additional questions were asked to further explore if participants understood the information that was provided to them and whether they used complex decision strategies (considering all attributes and making their choice based on trade-offs between all attributes) when making their choices. The majority of the participants seemed to have understood the provided information about the choice tasks, the attributes and the levels. They used complex decision strategies and are therefore capable of adequately completing a DCE. However, based on participants' age, educational level and health literacy, additional actions should be undertaken to ensure understanding of the choice tasks and completing the DCE as presumed. This third methodological study is described in **Chapter 10**.

In the general discussion (**Chapter 11**), the main findings of the previous chapters are integrated and further discussed alongside the methodological considerations. Directions for future research are formulated. The chapter concludes with the following recommendations:

RECOMMENDATIONS FOR HEALTH POLICY

1. During future pandemics, responsible authorities should align with important stakeholders in a country and communicate in a coordinated manner.
2. Adapt pandemic preparedness plans and communication strategies to local circumstances.
3. A premium discount of up to €15/month on restricted health plans might motivate especially younger, healthier and less wealthy consumers to choose these plans.

4. Policymakers who aim to expand the use of personal health records (PHRs) in the Netherlands will be most successful when healthcare providers or independent organizations facilitate the storage of PHR data, while refraining from including market parties in the storage of such data.
5. Target PHR information campaigns to chronically ill people.

RECOMMENDATIONS FOR DCE RESEARCHERS

1. Use qualitative methods (such as focus group discussions) to inform attribute and level selection, and report more details on this selection process.
2. Follow progress in methodology in research fields that have a longer tradition of using the DCE methodology.
3. Involve policy makers, or other end users of the DCE results, when designing the DCE. In addition, approach policy makers actively to present research findings and to discuss the implications for health policy. Consider publishing the DCE results in journals that are easily accessible for policy makers.
4. Use “opt-out” instead of “neither” in unforced DCEs.
5. Include a “status quo” alternative in addition to “opt-out” in markets where a status quo exists.
6. We find no indication that online surveys yield inferior results compared to paper based surveys. However, especially in online DCEs, researchers might want to include fewer choice sets per respondent to avoid fatigue or boredom of respondents.
7. Include age, education and health literacy measures in the DCE survey as well as in the analysis of DCE data. Until options to correct DCE responses for possible differences in demographic characteristics become common practice, researchers should at least describe these measures in their population and explain their possible effects on the results.

SAMENVATTING

Patiëntgerichtheid, gedeelde besluitvorming en zelfmanagement zijn fundamentele elementen van het huidige gezondheidsbeleid in Nederland. De betrokkenheid van burgers bij besluitvorming over gezondheidsbeleid wordt steeds meer gestimuleerd. Door het publiek te betrekken zal het beleid beter afgestemd zijn op publieke voorkeuren. Voorkeuren kunnen een impact hebben op de acceptatie, tevredenheid en uitkomsten van producten, diensten of interventies.

Hoewel tal van onderzoekstechnieken beschikbaar zijn om publieke voorkeuren voor gezondheid-gerelateerde onderwerpen te meten, gaat dit proefschrift specifiek over discrete keuze experimenten (DCE's). In een DCE vragenlijst krijgen respondenten een serie hypothetische keuzesituaties (ook wel keuzesets genoemd) voorgelegd, die elk bestaan uit twee of meer alternatieven van een product, dienst of interventie. De alternatieven worden beschreven aan de hand van hun kenmerken (ook wel attributen genoemd). De gepresenteerde alternatieven verschillen van elkaar door een systematische variatie in de niveaus van deze kenmerken (ook wel attribuut levels genoemd). Het woord 'discrete' geeft aan dat de respondenten slechts kunnen kiezen voor een van de alternatieven. De keuzes die respondenten maken worden geanalyseerd en geven inzicht in welke kenmerken het belangrijkst zijn en welke afwegingen gemaakt worden. Ook kan de deelnamebereidheid of het marktaandeel geschat worden. In het inleidende hoofdstuk (**Hoofdstuk 1**) worden relevante DCE begrippen geïntroduceerd.

DCE's zijn afkomstig uit de wiskundige psychologie en worden standaard toegepast in de marketing, en de transport- en milieueconomie. In 1990 werd de eerste DCE in de gezondheidseconomie gepubliceerd, en sindsdien neemt de hoeveelheid wetenschappelijke literatuur snel toe. Het doel van dit proefschrift is een bijdrage te leveren aan dit groeiende onderzoeksgebied, door onderzoeksvragen te beantwoorden die betrekking hebben op het verleden (deel I), het heden (deel II) en de toekomst (deel III) van de DCE methodologie. De onderzoeksvragen en de inhoud van dit proefschrift worden beschreven in **Hoofdstuk 1**.

Deel I: Het verleden - review van de literatuur

Hoofdstuk 2 geeft een overzicht van de tussen 2009 en 2012 gepubliceerde gezondheid-gerelateerde DCE's. Ook worden methodologische ontwikkelingen sinds twee eerder gepubliceerde reviews (over de periodes 1990-2000 en 2001-2008) beschreven in dit hoofdstuk. We identificeerden 179 gezondheid-gerelateerde DCE's in PubMed (een database met referenties van wetenschappelijke artikelen). We vonden dat kwalitatieve methoden (zoals focusgroepen) minder vaak werden gebruikt om attributen te selecteren in vergelijking met de vorige periodes. Daarnaast vonden we een toename in het gebruik van zogenaamde 'D-efficient' designs en meer flexibele econometrische methoden, waaronder

'mixed logit' en 'latent class' modellen. DCE data wordt in toenemende mate digitaal en/of via Internet panels verzameld.

Deel II: Het heden - drie state-of-the-art toepassingen
Het tweede deel van dit proefschrift beschrijft een kwalitatieve studie en drie DCE toepassingen (pandemische vaccinaties, zorgverzekeringen en persoonlijk gezondheidsdossiers).

De eerste drie hoofstukken van dit deel van het proefschrift zijn geschreven naar aanleiding van het project 'Effective Communication in Outbreak Management: development of an evidence-based tool for Europe (ECOM)'. **Hoofdstuk 3** beschrijft een kwalitatieve verkenning van publieke opinies en attitudes over toekomstige pandemieën en vaccinaties in drie Europese landen (Nederland, Zweden en Polen). In elk land werden twee focusgroepen met 6-9 deelnemers per focusgroep gehouden. Deelnemers gaven aan hun vaccinatiebeslissingen te baseren op afwegingen tussen de verwachte voor- en nadelen van een vaccin (zoals werkzaamheid en bijwerkingen). In die beslissing wegen zij ook de ernst van de nieuwe pandemie mee. Behalve degenen met één of meer chronische ziekten, verwachten deelnemers niet dat zij snel ziek zullen worden tijdens een toekomstige pandemie. Alle deelnemers gaven aan bezorgd te zijn over de veiligheid van de vaccins, vanwege de beperkte tijd om vaccines te produceren en te testen in de acute situatie van een pandemie. Belangrijke verschillen tussen de deelnemers uit verschillende landen werden gevonden voor eerdere vaccinatie ervaringen, de invloed van sociale normen op gedrag en het vertrouwen in de gezondheidsautoriteiten. De bevindingen van deze studie dienden als input voor de selectie van de attributen en attribuut levels voor de eerste DCE toepassing over voorkeuren voor pandemische vaccinaties beschreven in **Hoofdstuk 4** en **5**. Een vergelijking van voorkeuren voor pandemische vaccinaties tussen vier Europese landen (Nederland, Spanje, Zweden en Polen) wordt beschreven in **Hoofdstuk 4**. Leden van Internet panels, landelijk representatief op basis van leeftijd, geslacht, opleidingsniveau en regio (N=2068) vulden in 2013 een online DCE vragenlijst. In het geval van een ernstige toekomstige pandemie was de werkzaamheid van een vaccin het belangrijkste kenmerk dat vaccinatie voorkeuren in vier landen bepaald. Dit werd gevolgd door het advies over het vaccin. Het advies van familie en/of vrienden en het advies van artsen beïnvloedde vaccinatie voorkeuren van Zweedse respondenten het meeste. Voor Poolse en Spaanse respondenten was vooral het advies van de (internationale) gezondheidsinstanties van belang. Ongeacht het pandemische scenario of de kenmerken van een vaccinatieprogramma was de geschatte vaccinatiegraad het laagst in Zweden en het hoogst in Polen. De DCE data van **Hoofdstuk 4** werd verder gebruikt om verschillen in voorkeuren voor pandemische vaccinaties binnen Nederland (**Hoofdstuk 5**) te bestuderen. Vrouwen en respondenten die verklaren nooit vóór vaccinatie te zijn, maakten andere afwegingen bij de beslissing of zij zich al dan niet zouden laten vaccineren in het geval van een pandemie dan manen en respondenten die verklaren dat zij (mogelijk) bereid zijn zich te laten vaccineren.

Hoofdstuk 6 heeft betrekking op de tweede DCE toepassing van dit proefschrift, waarin de afwegingen die consumenten maken als zij kiezen voor de basiszorgverzekering in kaart werden gebracht. Ook werd onderzocht of, en welke verschillen er zijn in afwegingen op basis van leeftijd, gezondheidstoestand en inkomen. Een representatieve steekproef (N=533) van de Nederlandse volwassen bevolking voltooide de online DCE vragenlijst tijdens de tweemaandelijkse periode waarin van zorgverzekering gewisseld kon worden voor het jaar 2015. Het hebben van vrije artsenkeuze (en de bijbehorende vergoeding van zorgkosten) was veruit het meest beslissende kenmerk voor de keuze voor een zorgverzekering voor respondenten ouder dan 45 jaar, respondenten met één of meer chronische aandoeningen en respondenten met een bruto inkomen van meer dan €3000 per maand. Hoewel de vrije artsenkeuze (en bijbehorende vergoeding van zorgkosten) ook een belangrijk kenmerk was voor jongere en gezondere respondenten en respondenten met een lager inkomen, was de maandelijkse premie voor hen het belangrijkste voor hun keuze. Zorgverzekeringen van zorgverzekeraars die bij het inkopen van zorg focussen op kwaliteit van zorg werden verkozen boven verzekeringen van zorgverzekeraars met een focus op goedkopere aanbieders, hoewel de betalingsbereidheid laag was. De helft van de steekproef zou nooit kiezen voor een polis met een beperkte keuzevrijheid. Echter een premiekorting van maximaal €15 per maand op polissen met beperkt keuzevrijheid zal vooral jongere, gezondere en minder welvarende consumenten motiveren om zo een polis te kiezen.

Hoofdstuk 7 heeft betrekking op de derde DCE toepassing over voorkeuren voor persoonlijke gezondheidsdossiers (PGDs). Deze studie richt zich specifiek op personen die op dit moment geen PGD hebben. Een PGD is een digitaal hulpmiddel voor patiënten waarin informatie over gezondheid en ziekte kan worden verzameld, beheerd en gedeeld. De DCE vragenlijst werd in december 2015 en januari 2016 ingevuld door 1443 Internet panel leden van de Patiëntenfederatie NPCF. We identificeerden drie subgroepen met verschillende voorkeuren tussen de groepen: zij die gretig zijn, zij die terughoudend zijn en zij die weigeren. Respondenten met één of meer chronische ziekten hadden een hogere kans om tot de gretige subgroep te behoren. De faciliterende partij voor de opslag van PGD data was het meest doorslaggevende kenmerk voor zowel de gretige als de terughoudende subgroep, terwijl de kosten het meest bepalend waren voor de weigeraars. In alle drie de groepen waren zorgverleners en onafhankelijke organisaties de meest geprefereerde facilitators van dataopslag. De geschatte deelnamebereidheid van de gretige subgroep was altijd boven de 91%, de deelnamebereidheid van de terughoudende subgroep varieerde tussen de 4% en 68% en was afhankelijk van PHR kenmerken, en de deelnamebereidheid van de weigeraars was altijd onder de 9%.

Deel III: De toekomst - methodologische studies
Het derde deel van dit proefschrift beschrijft drie methodologische DCE studies.

In keuzesets wordt, naast de hypothetische alternatieven van producten, diensten of interventies, vaak een "opt-out" optie geboden. Dit betekent dat als je voor deze optie kiest, je het product niet wilt kopen in de gegeven situatie, of geen gebruik wilt maken van de interventie of dienst. Minder vaak wordt gekozen voor de optie "geen van beiden" (de optie om geen van beide alternatieven te kiezen) en een "status quo" alternatief (de optie om te kiezen voor het huidige alternatief in plaats van de hypothetische alternatieven). In **Hoofdstuk 8** hebben we de implicaties van het gebruik van deze verschillende formats getest. Deze methodologische studie was onderdeel van de DCE over het persoonlijk gezondheidsdossiers (**Hoofdstuk 7**). We toonden aan dat de keuze voor een bepaald format van invloed is op de resultaten en dat dit meer aandacht zou moeten krijgen van DCE onderzoekers. Wij adviseren om "opt-out" te gebruiken in plaats van "geen van beiden", omdat "geen van beiden" multi-interpretabel is voor respondenten. "Geen van beiden" wordt slechts zelden geïnterpreteerd als een echte "opt-out", terwijl dit door onderzoekers wel zo gemodelleerd wordt. Daarnaast adviseren wij om zowel een "opt-out" als een status quo op te nemen in keuzesets voor markten waarin een status quo bestaat.

DCE data wordt steeds vaker verzameld via digitale vragenlijsten en/of Internet panels. In de studie beschreven in **Hoofdstuk 9** onderzochten we of de manier van dataverzameling (via papier of online) van invloed is op de uitkomsten van een DCE. Deze methodologische studie was onderdeel van de DCE over zorgverzekeringen (**Hoofdstuk 6**). We vonden geen aanwijzingen dat dataverzameling via online vragenlijsten minder goede resultaten zou opleveren dan papieren vragenlijsten, terwijl de prijs per respondent over het algemeen lager is voor online vragenlijsten.

Om de juiste conclusies te trekken op basis van DCE data, is het essentieel dat onderzoekers begrijpen hoe deelnemers de keuzesets in een DCE invullen. Om dit te onderzoeken hebben we 70 gestructureerde interviews gehouden. Deelnemers werden eerst gevraagd om vier keuzesets hardop in te vullen. Vervolgens werden aanvullende vragen gesteld om verder te verkennen of de deelnemers de verstrekte informatie begrepen en of ze alle attributen tegen elkaar afwogen bij het maken van hun keuzes. De meerderheid van de deelnemers leek de verstrekte informatie over de keuzesets, de attributen en de attribuutlevels te hebben begrepen en woog de meerderheid van de attributen tegen elkaar af. Echter, op basis van leeftijd, opleidingsniveau en gezondheidsvaardigheden zullen aanvullende maatregelen genomen moeten worden om te zorgen dat keuzesets worden ingevuld zoals verondersteld wordt door onderzoekers. Deze derde methodologische studie is beschreven in **Hoofdstuk 10**.

In het laatste hoofdstuk van dit proefschrift (**Hoofdstuk 11**) worden de belangrijkste bevindingen van de voorgaande hoofdstukken geïntegreerd en verder bediscussieerd. Methodologische overwegingen en aanknopingspunten voor toekomstig onderzoek worden benoemd. Het hoofdstuk sluit af met de volgende aanbevelingen:

AANBEVELINGEN VOOR GEZONDHEIDSBELEID

1. Tijdens toekomstige pandemieën is het wenselijk dat verantwoordelijke autoriteiten samenwerken met belangrijke stakeholders in een land, en op een gecoördineerde manier communiceren.
2. Pas pandemiedraaiboeken en communicatiestrategieën aan nationale omstandigheden aan.
3. Een premiekorting van maximaal €15 per maand op polissen met beperkt keuzevrijheid zal vooral jongere, gezondere en minder welvarende consumenten motiveren om zo een polis te kiezen.
4. Beleidsmakers die als doel hebben om het gebruik van persoonlijke gezondheidsdossiers (PGD) in Nederland te vergroten zullen het meest succesvol zijn als zorgverleners of onafhankelijke organisaties de dataopslag van PGDs faciliteren, in plaats van commerciële bedrijven.
5. Focus in PGD voorlichtingscampagnes op chronisch zieken.

AANBEVELINGEN VOOR ONDERZOEKERS DIE DE DCE METHODOLOGIE GEBRUIKEN

1. Gebruik kwalitatieve technieken (bijvoorbeeld focusgroepen) om attributen en attribuut levels te selecteren en rapporteer meer details over dit selectieproces.
2. Volg de vooruitgang in de methodologie in onderzoeksvelden die de DCE methodologie al langer gebruiken.
3. Betrek beleidsmakers en andere eindgebruikers van de DCE resultaten bij het opzetten van de DCE. Benader daarnaast actief beleidsmakers om onderzoeksresultaten te presenteren en de gevolgen voor beleid te bespreken. Overweeg de DCE resultaten te publiceren in tijdschriften die gemakkelijk toegankelijk zijn voor beleidsmakers.
4. Gebruik een “opt-out” optie in plaats van een “geen van beiden” optie.
5. Neem een “status quo” alternatief als aanvulling op de “opt-out” optie op in keuzesets voor markten waarin een status quo bestaat.
6. We vinden geen aanwijzingen dat online vragenlijsten minder goede resultaten opleveren in vergelijking met papieren vragenlijsten. Echter, vooral als DCE data online wordt verzameld, zouden onderzoekers kunnen overwegen om minder keuzesets per

respondent in de DCE op te nemen om vermoeidheid of verveling van de respondenten te voorkomen.

7. Neem vragen op over leeftijd, opleiding en gezondheidsvaardigheden in de DCE vragenlijst en in de analyse van de DCE data. Tot mogelijkheden om DCE antwoorden te corrigeren voor mogelijke verschillen in demografische kenmerken standaard toegepast worden, moeten de onderzoekers in ieder geval de antwoorden op deze vragen rapporteren en de mogelijke effecten op de resultaten beschrijven.

Dankwoord

List of publications

PhD portfolio

About the author

DANKWOORD

Ruim vier jaar geleden stond ik voor de (discrete) keuze: kies ik voor dit promotieonderzoek, of toch voor iets anders? Degenen die mij kennen weten dat ik niet altijd een ster ben in het maken van keuzes - het liefst wil en doe ik alles. Achteraf gezien paste dit promotieonderzoek daarom ook zo goed bij mij: de afgelopen vier jaar waren afwisselend, leerzaam en vol met mooie kansen. Op deze plaats wil ik graag iedereen bedanken die, ieder op zijn of haar eigen manier, heeft bijgedragen aan de totstandkoming van dit proefschrift.

Mijn promotor en copromotoren: Ewout Steyerberg, Esther de Bekker-Grob en Mattijs Lambooij. Ewout, bedankt voor je snelle reacties, je kritische blik en je toegankelijkheid. Tijdens onze afspraken kwam onze gedeelde passie voor hardlopen altijd eerst aan bod. De inhoudelijke en wetenschappelijke discussies die daarna volgden hebben absoluut bijgedragen aan mijn wetenschappelijke ontwikkeling. Esther, ik waardeer het vertrouwen dat je in mij had en de zelfstandigheid die je mij gaf. Ik wens je alvast een plezierige en nuttige tijd in Australië toe. Mattijs, bedankt voor je betrokkenheid, je ideeën en je creativiteit. Ik heb met plezier met je samengewerkt.

Ardine de Wit en Ida Korfage, ondanks dat jullie niet officieel als copromotoren in dit proefschrift genoemd mochten worden, zie ik jullie weldegelijk als copromotor. Ardine, altijd waren je reacties op mijn stukken opbouwend en zeer waardevol, zowel inhoudelijk als tekstueel. Ik hoop in een latere fase van mijn carrière op een vergelijkbare manier te kunnen bijdragen aan artikelen. Ida, je ving mij op mijn eerste werkdag op en ook daarna was je er altijd voor me. Onze samenwerking voelde als vanzelfsprekend. Bedankt voor je oprechte geïnteresseerdheid, zowel werkinhoudelijk als privé.

Ik wil de leden van de kleine en grote promotiecommissie bedanken voor de tijd die aan mijn proefschrift is besteed en voor het opponeren tijdens de verdediging.

De artikelen in dit proefschrift zouden er niet gekomen zijn zonder samenwerking met andere onderzoekers binnen én buiten Nederland. Ik wil alle coauteurs bedanken voor hun waardevolle inbreng. Graag noem ik een aantal mensen in het bijzonder: Jan Hendrik Richardus en Helene Voeten, bedankt voor jullie inzet om het ECOM project tot een succesvol einde te brengen. Helene, ik zal je altijd dankbaar blijven voor de beste boekentip ooit! Thanks also to all other ECOM project members; it was a pleasure to be part of this very interesting project. I really enjoyed working with all of you. Michiel Bliemer, Bas Donkers en Marcel Jonker, bedankt dat we indien nodig een beroep op jullie DCE expertise konden doen. Erik Schut, Marco Varkevisser en Arthur Hayen, jullie input betreffende de zorgverzekeringen DCE was zeer waardevol. Line Pedersen og Dorte Gyrd-Hansen, tak for muligheden for at komme til jeres universitet. Tak også for den tid I har investeret i mig og i projektet. Jeg har

virkelig lært en masse af jer. Det var en berigelse af min PhD! Jeg håber, at vi vil samarbejde igen i fremtiden. Line, jeg takker dig og Martin også for jeres gæstfrihed og hyggelig tid i Odense: jeg er glad for jeg mødte dig. De artikelen in dit proefschrift zouden er ook niet zijn gekomen zonder de deelnemers aan de interviews, focusgroepen en vragenlijsten. Bedankt!

Het was een voorrecht om zowel bij het Erasmus MC als het RIVM werkzaam te zijn en dus ook dubbel zoveel fijne collega's te hebben. Ik wil jullie allemaal, maar in het bijzonder mijn (buur-, en flex-) kamergenoten, NIHES maatjes, JVO'ers, CMB'ers, KZG'ers, en lunchwandelaars, bedanken voor de samenwerking en collegialiteit. Ik wil de dames van beide secretariaten en de ICT'ers bedanken voor hun hulp en ondersteuning. Ik wens iedereen die nog met een promotieonderzoek bezig is succes met het vervolg hiervan. Esther H., ik ga onze Spotify-sessies missen! Yesim, het was fijn om onze ervaringen met elkaar te kunnen delen. Overigens ben ik blij voor je dat je het naar je zin hebt in de dermatologie. Jorien, ik denk met plezier terug aan onze tijd samen op het RIVM. Wat hebben we gelachen tijdens het schrijven van ons artikel. Juist door onze verschillende manieren van werken vulden we elkaar goed aan.

Fenna, Kerstin, en Rienke, we zijn ondertussen geen collega's meer, en daarom is dit zo'n mooie overgang naar de vrienden alinea. Ik ben zo blij dat we elkaar hebben leren kennen. Met jullie is het altijd gezellig waar we ook zijn en wat we ook doen: in de trein van 7.17u, tijdens onze hardlooprondjes, als we koffiedrinken met lekkere baksels (ik ben een bofkont!) en ook tijdens het kaarten. Bedankt voor jullie vriendschap. Ik wil ook mijn andere vrienden en vriendinnen bedanken voor hun betrokkenheid, waarvan vier in het bijzonder. Anne, het briefje "Veel plezier en succes op je eerste werkdag. Wel drie treinen eerder nemen hè?' dat je schreef heb ik nog steeds! Het was leuk dat we huisgenoten waren in die tijd en onze eerste werkervaringen met elkaar konden delen, hoe anders die ervaringen ook waren. Willemien, mede-ranger! Takk voor de afleiding en gezelligheid die onze reizen altijd boden. Binnenkort maar weer eens een reisje plannen? Marjolein, ik was altijd welkom om bij jullie te eten als ik in Rotterdam had gewerkt. Dat was leuk, thanks. Het werd iets minder praktisch sinds jullie zijn verhuisd, maar ook nu weet ik dat ik altijd welkom ben. Marlies, ik vind het zo fijn dat we al zo lang vriendinnen zijn en hoop dat we dat nog véél langer blijven. Ik voel mij ontzettend op mijn gemak bij jou en ik ben daarom ook heel blij dat je mijn paranimf wilt zijn. Dit is dan weer een mooie overgang naar de familie alinea, want je zult altijd 'La Viana' blijven, ook al ben je al weer een paar jaar 'La de Vrij'.

Lieve familie en schoonfamilie: bedankt voor jullie interesse de afgelopen jaren. Oma John & John, leuk dat jullie mijn artikelen lezen! Jammer dat opa Kippen niet meer met ons mee kan vieren dat mijn 'studie' klaar is en ik een 'echte' baan heb. Lieve Laszlo en Samantha: ja nu ben ik dan eindelijk PhDomino! Ook door jullie mede mogelijk gemaakt; tusentakk voor jullie input. Ik word er echt heel blij van hoe gezellig we het met z'n vieren hebben. Fijn dat

je mijn paranimf wilt zijn lieve broer. Lieve Mam en Tino, wat leven jullie altijd met mij mee. Mam, dat doe jij op belangrijke momenten, maar juist ook zomaar, wanneer je mij een fijne werkdag wenst als je weet (of denkt, of hoopt ;)) dat ik in de trein zit of net op m'n werk ben aangekomen. Dat doet mij goed. Tekenend voor je betrokkenheid is ook dat je inmiddels regelmatig allerlei vragenlijsten invult (en de mijne dan, natuurlijk, toch vaak beter in elkaar vindt zitten). Lieve Pepz, wat was het fijn om allebei in Rotterdam te werken! Altijd maak je tijd voor mij vrij om te lunchen of om onder het genot van een troebel appelsapje in de Rotterdamse Schouwburg de dag te bespreken. Gelukkig is mijn nieuwe baan ook weer een beetje bij jou in de buurt. Bedankt ook voor je originele ideeën over de voorkant van mijn proefschrift.

Lieve Len, de laatste woorden zijn voor jou. Thanks voor je support en het aanhoren van al mijn verhalen. Ik kan nog veel van jouw positiviteit en geduld leren. Ik kijk uit naar onze toekomst samen: we zijn een goed duo!

Domino

LIST OF PUBLICATIONS

By July 22, 2016

2014

†Determann D, Korfage IJ, Lambooij MS, Bliemer MC, Richardus JH, Steyerberg EW, de Bekker-Grob EW. Acceptance of vaccinations in pandemic outbreaks: a discrete choice experiment. PLoS ONE. 2014;9(7):e102505. doi:10.1371/journal.pone.0102505

†Clark M, Determann D, Petrou S, Moro D, de Bekker-Grob EW. Discrete choice experiments in health economics: a review of the literature. Pharmacoeconomics. 2014;32(9):883-902. doi:10.1007/s40273-014-0170-x

2015

Determann D, Hameli L, Klik M, Roos, AK, de Ruiter H, van der Silk E, Spoon EQW, Korfage IJ. Vaccinatiebereidheid van toekomstige artsen tijdens een volgende Influenza-pandemie. Infectieziekten bulletin. 2015;26(3).

2016

†Determann D, de Bekker-Grob EW, French J, Voeten HA, Richardus JH, Das E, Korfage IJ. Future pandemics and vaccinations: public opinion and attitudes across three European countries. Vaccine. 2016;34(6):803-808. doi:10.1016/j.vaccine.2015.12.035

†Veldwijk J*, Determann D*, Lambooij MS, van Til JA, Korfage IJ, de Bekker-Grob EW, de Wit GA. Exploring how individuals complete the choice tasks in a discrete choice experiment: an interview study. BMC Medical Research Methodology 2016;16(45). doi: 10.1186/s12874-016-0140-4

†Determann D, Korfage IJ, Fagerlin A, Steyerberg EW, Bliemer MC, Voeten HA, Richardus JH, Lambooij MS, de Bekker-Grob EW. Public preferences for vaccination programmes during pandemics caused by pathogens transmitted through respiratory droplets - a discrete choice experiment in four European countries, 2013. Eurosurveillance. 2016;21(22):pii=30247 doi:10.2807/1560-7917.ES.2016.21.22.30247

†Determann D, Lambooij MS, de Bekker-Grob EW, Hayen AP, Varkevisser M, Schut FT, de Wit GA. What health plans do people prefer? The trade-off between premium and provider choice. Accepted for publication in Social Science & Medicine. July 2016

* Authors contributed equally to the manuscript
† In this thesis

PHD PORTFOLIO

Summary of PhD training and teaching

Name:	Domino Determann
Erasmus MC department:	Public Health
Research school:	Netherlands Institute for Health Sciences (NIHES)
PhD period:	2012-2016
Promotor:	Prof. Dr. EW Steyerberg
Copromotors:	Dr. EW de Bekker-Grob (Erasmus MC), Dr. MS Lambooij (RIVM)

	Year	Workload (ECTS)
1. PhD Training		
General academic skills		
- Research integrity for PhD students, Erasmus MC	2014	0.3
- Time and project management course for PhD students, Erasmus MC	2012	0.6
In-depth courses		
- Health economics introductory course, RIVM	2015	2.0
- Master of Science in Health Sciences: specialization clinical epidemiology, NIHES	2012-2015	70
Discrete choice experiment courses		
- Choice modelling course, Institute for Transport Studies, University of Leeds, United Kingdom	2014	1.5
- Workshop 'Using discrete choice experiments in health economics; theoretical and practical issues', Health Economics Research Unit (HERU), University of Aberdeen, United Kingdom	2013	0.9
National and international conferences		
- European Scientific Conference on Applied Infectious Disease Epidemiology (ESCAIDE), Stockholm, Sweden: poster and oral presentation	2015	0.9
- Annual European Congress of the International Society for Pharmacoeconomics and Outcomes Research (ISPOR), Amsterdam: poster presentation	2014	0.9
- International Academy of Health Preferences Research (IAHPR), Amsterdam	2014	0.3
- Lowlands Health Economics Study Group conference (LolaHESG), Oostvoorne	2014	0.6
- Biennial European Conference of the Society of Medical Decision Making (SMDM), Antwerp, Belgium: oral presentation	2014	0.6
- European Congress of Clinical Microbiology and Infectious Diseases (ECCMID), Barcelona, Spain: oral presentation	2014	1.2
- Nederlands Congres voor Volksgezondheid (NCVGZ), Rotterdam: poster presentation	2014	0.6
- Lowlands Health Economics Study Group conference (LolaHESG), Nunspeet	2013	0.6

Seminars and workshops		
- 'A free lunch' health economics meetings, RIVM	2014-2016	1.0
- Seminars department of Public Health, Erasmus MC	2012-2016	1.0
- Meetings Medical Decision Making, department of Public Health, Erasmus MC	2012-2016	1.0
- Pre-ESCAIDE final symposium ECOM project, Stockholm, Sweden: oral presentation	2015	0.3
- Using stated preference methods in health and public health, RIVM: oral presentation	2015	0.6
- Minisymposium 'Risicocommunicatie bij opkomende infectieziekten', Erasmus MC	2014	0.2
Other presentations		
- Oral presentations within department of Public Health, Erasmus MC	2013-2016	0.6
- Oral presentation at Epidemiology meeting, RIVM	2015	0.2
- Oral presentation at brown-bag seminar, COHERE, Odense, Denmark	2015	0.2
- Oral presentation at internal seminar applied economics, COHERE, Odense, Denmark	2015	0.2
- 'Eenduidige boodschap cruciaal in strijd tegen toekomstige pandemieën' (2014) and 'NL-er wil bij elke arts terecht kunnen; kwaliteit van zorg is ondergeschikt in keuze voor zorgverzekering' (2015), oral presentations ('half tientjes'), RIVM	2014-2015	0.2
- Oral presentations at ECOM project meetings in Munster (2013), Rotterdam (2014), London (2014), Hamburg (2015)	2013-2015	5.0
Didactic skills		
- 'Teach the teachers' didactic training Erasmus MC, part-qualification teaching	2015	0.6
- 'VO Lifestyle counselling and primary prevention' didactic training Erasmus MC, part-qualification teaching	2013	0.2
2. Teaching activities		
- Lecturer 'Patient preferences in the delivery of healthcare', Master HEPL, EUR	2016	0.3
- Lecturer VO Vaccination training (theme 3.C.3)	2015-2016	1.0
- Correcting QALY exam question (theme 3.C.1)	2015	0.3
- Correcting Bachelor essays (3rd year medical students)	2014	2.0
- Correcting discourses (Medical Research and Society module, 2nd year medical students)	2014	1.0
- Assistant teacher 'Patient preferences in the delivery of healthcare', Master HEPL, EUR	2014	0.3
- Wetenschapsknooppunt EUR: introduction to science at primary and secondary schools	2013-2014	0.3
- Supervisor Community projects (theme 3.C.4)	2012-2014	2.0
- Lecturer VO Lifestyle counselling and primary prevention (theme 3.C.2)	2013	0.6
3. Other activities		
- Peer reviewer for scientific journals (BMC Health Services Research, Health Policy and Planning, BMJ Open)	2014-2016	N/A
- Member of the junior representatives consultation, department of Public Health, Erasmus MC	2013-2016	N/A

- Visiting PhD student, Centre of Health Economics Research (COHERE), University of Southern Denmark, Odense, Denmark	2015	N/A
- Advisory board 'MMRV medical education communication strategy and curriculum development', Glaxo Smith Kline, London, United Kingdom	2014	0.9
- Career orientation training, Erasmus MC	2014	0.6
- PhD day, Erasmus MC	2012	0.3
4. Awards		
- Epiet Alumni Network (EAN) prize for best poster presentation at ESCAIDE 2015	2015	N/A

ABOUT THE AUTHOR

Domino Determann was born on the 8th of May 1987 in Rotterdam, the Netherlands. In 2005, she finished secondary school (Antonius College, Gouda). In this year, she also started studying Medicine at the University of Utrecht. In addition to the standard curriculum, Domino followed a course on 'Economics and financing of healthcare systems' at the Institute of Health Policy & Management (iBMG, Erasmus University, Rotterdam), and did an elective 'Health policy & management at the Dutch Healthcare Authority (NZa). Furthermore, she went to Trondheim, Norway, to do her Ear-Nose-Throat internship. She obtained her doctoral degree in December 2011. Subsequently Domino started her PhD trajectory on discrete choice experiments in April 2012, jointly at the department of Public Health (Erasmus MC) and at the Centre for Nutrition, Prevention and Health Services (National Institute for Public Health and the Environment, RIVM). She conducted part of her PhD research in collaboration with the researchers of the Centre of Health Economics Research (COHERE) of the University of Southern Denmark in Odense. She hereto worked and lived in Odense for a period of three months in 2015. As part of her PhD trajectory, Domino obtained a master's degree (with an average grade of 8.4) in Health Sciences, specialisation Clinical Epidemiology, in 2015 at the Netherlands Institute for Health Sciences (NIHES). As of June 2016, Domino is working as health consultant at Ecorys in Rotterdam.

* The quintessence of a discrete choice experiment