Do people's choices reflect their opinions? The association between Q-methodology and Participatory Value Evaluation in the context of disinvestment of healthcare in The Netherlands

Elisa de Weerd (505655ew)

NIHES Research Master Health Sciences, Health Economic Analysis

Erasmus School of Health Policy & Management, Erasmus University Rotterdam

RIVM, National Institute for Public Health and the Environment

Supervised by prof. dr. Job van Exel and Adriënne Rotteveel, MSc

3rd of July, 2020

Abstract

There are many ways to measure preferences. The question remains to what extent these methods are associated and yield comparable results. This study explores the relationship between findings of a recent Q-methodology study and a Participatory Value Evaluation (PVE) in the context of active disinvestment of healthcare in The Netherlands. Using a subgroup analysis, we found that there are only limited, and illogical differences in PVE findings between people with different viewpoints. This can be the case because opinions and preferences in a choice experiment are largely unrelated, because the differences are driven by other factors than opinions, or because the measurement of preferences in the studies was not specific enough to distinguish differences between people who have different opinions. Further analysis could explore the relationship between Q-methodology and PVE using interactions or with a latent class analysis. In addition, we also explored the distribution of viewpoints in the Dutch population. We found that the majority of Dutch citizens think that decision-making should be transparent and consistent. Furthermore, all four viewpoints found in the analysed Q-methodology study are reflected in the Dutch population.

Keywords: Participatory Value Evaluation, Q-methodology, disinvestment, healthcare, preferences, opinions, choices

Project group

This research is part of an internship at the Strategic Programme RIVM (National Institute for Public Health and the Environment) (SPR) 'QALYs foregone' and 'CIVILIAN'. The thesis was carried out under supervision of prof. dr. Job van Exel and PhD candidate Adriënne Rotteveel. They are both active in the Health Economics unit of the Erasmus School of Health Policy & Management (ESHPM). Adriënne Rotteveel carries out her PhD project at the RIVM, Centre for Nutrition, Prevention and Health Services, at the department of Quality of Care and Health Economics.

Acknowledgements

I would like to express my great appreciation to all who have contributed to this study. Adriënne Rotteveel has introduced me to a very interesting research project, and provided valuable feedback, together with Job van Exel, along the way. Your eagerness to learn from the research findings further increased my interest in the topic. Advice and explanations on the PVE design given by everyone in the CIVILIAN Project Group has also been greatly appreciated. I am particularly grateful for the enthusiasm and assistance with the PVE analysis of Niek Mouter, José Ignacio Hernández and Eelco Over. The students from the TU Delft PVE Thesis Circle also provided valuable insights on my topic, as well as sparked my enthusiasm for preference studies in other contexts. I am happy that I got the chance to do this study as part of an internship at the National Institute of Public Health and the Environment. Even though the COVID-19 crisis has led me to work from home most of the time, I learned a lot from the department and am glad that I had the opportunity to experience working at this institute.

Table of Contents

Introduction

Measuring preferences regarding healthcare goods and services is challenging. Since the production and distribution of healthcare goods and services are usually not determined by prices, data on revealed preference is unavailable. The evaluation of actual choice decisions in healthcare is thus often infeasible and other methods should be used to elicit preferences. However, it is particularly in healthcare that measuring preferences is important. As a result of expensive technologies, aging populations and capacity problems in care staffing, healthcare costs are rising.[1-3] This creates pressure on public expenditures. Rising public healthcare expenditures could thus crowd out other public expenditures. This means that trade-offs need to be made.

Decision-making on diminishing healthcare expenditure is, however, societally sensitive. The importance of careful considerations and legitimate decision-making is therefore high. In order to prevent societal discontent or protest, it is important that the public opinion is taken into account.

Public preferences in healthcare can be measured in different ways, which all have their advantages and disadvantages. In the last decades, several new methods have been developed which aim to overcome part of the challenges of measuring preferences. The question remains to what extent these methods are associated and yield comparable results.

Recently, two preference studies were conducted that research people's preferences regarding disinvestment of healthcare, which is one of the options to diminish the rise of healthcare expenditure. In 2019, a Q-methodology study was conducted by Rotteveel, Reckers-Droog, Lambooij, de Wit, van Exel [4]. They found that four societal views exist in the Dutch population regarding active disinvestment of healthcare. In the Spring of 2020, a Participatory Value Evaluation (PVE) was conducted that studies the decisions that people made when they were asked to reduce healthcare costs by cutting reimbursement worth 100 million Euros. By

analysing choice responses, the relative importance of each treatment characteristic could be inferred. Next to completing the choice task, the participants of the PVE. were asked about the extent to which they agree to the societal views found in Rotteveel, Reckers-Droog, Lambooij, de Wit, van Exel [4]. In this study, we will assess whether associations exist between the extent to which someone agrees with a societal view, and their decisions in the PVE.

This is done by answering the research question: *To what extent are opinions and preferences in a choice experiment associated?* Specifically, we study to what extent the societal views on disinvestment of healthcare interventions discerned by Rotteveel, Reckers-Droog, Lambooij, de Wit, van Exel [4] are associated with decision making patterns found in the PVE on disinvestment. In addition, the representative sample used in the PVE study also allows us to explore the distribution of viewpoints found in the Q-methodology study in the Dutch population. Therefore, the second research question is: *What is the distribution of viewpoints regarding active disinvestment in the Dutch population?*

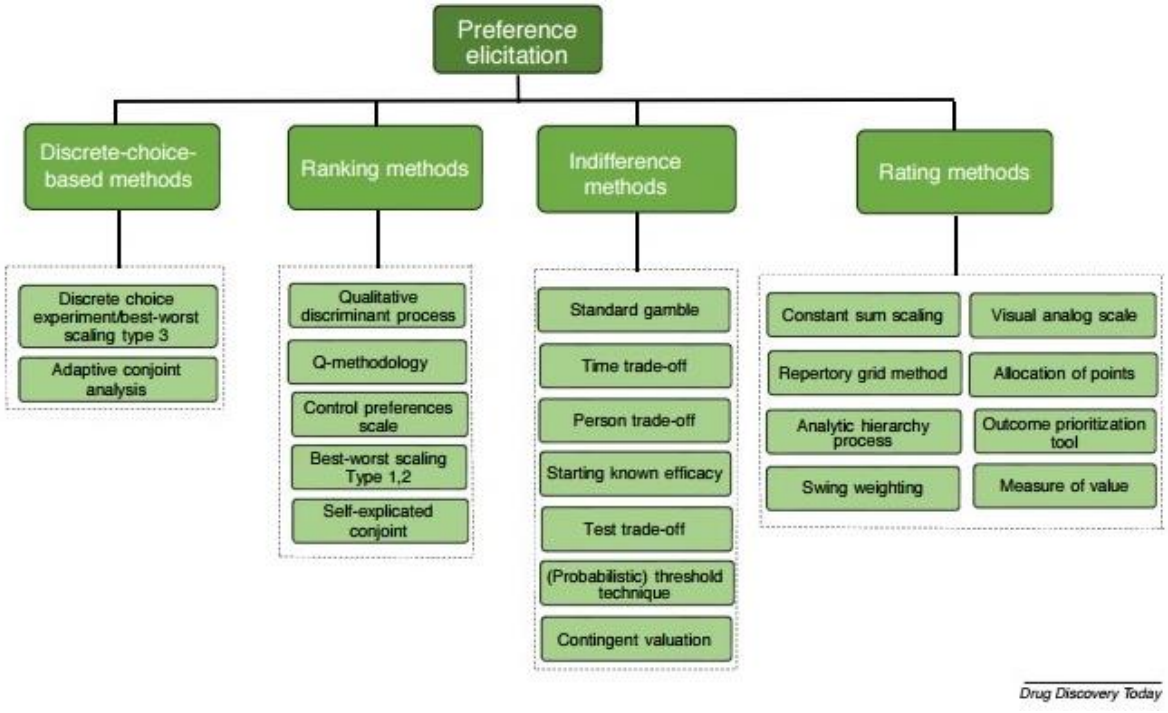
Literature

Preferences can be measured in two ways: as stated preferences and as revealed preferences. Revealed preference approaches measure consumer behaviour to estimate the utility the consumer gains from a certain product or service. An example to elicit this is the hedonic pricing method, which is a model that relies on the assumption that prices are determined both by internal as well as external factors. Criticism on this approach is that it relies too much on assumptions such as continuity of preference over time. In public healthcare contexts, revealed preference methods are often infeasible to use because of the lack of prices.

Stated preference techniques ask people what they find important in a certain decision; what economic value they attach to an attribute (characteristic). Examples of stated preference techniques are choice modelling and contingent valuation. According to Soekhai, Whichello,

Levitan, Veldwijk, Pinto, Donkers, Huys, van Overbeeke, Juhaeri, de Bekker-Grob [5], preference elicitation of stated preference can be divided into four categories: discrete-choice based methods, ranking methods, indifference methods and rating methods.

Figure 1: Grouping of preference elicitation (quantitative) methods into four groups (source: Soekhai, Whichello, Levitan, Veldwijk, Pinto, Donkers, Huys, van Overbeeke, Juhaeri, de Bekker-Grob [5])



In choice-based methods, respondents are asked to choose their preferred good out of a choice set. An example of this is a Discrete Choice Experiment (DCE). The options in the choice sets are designed using a set of attributes, each with several levels. This results in trade-offs between attributes. If a monetary attribute is included, monetary payments per attribute can be elicited. The method is well suited to situations with competing options or priorities, but the information requirements are generally greater than what is needed to design other preference methods.

In studies which use a ranking method, respondents are asked to rank certain statements or characteristics in order to explore what relative importance people attach to certain statements or characteristics over the others. An example of a ranking method is Q-methodology, which combines qualitative methods with quantitative methods. It is an

adaptation of Spearman's method of factor analysis.[6] Participants of a Q-methodology study are presented with a set of statements and are asked to respond to each statement with the extent to which they agree with them. It is meant to reveal patterns of association between a series of measurable variables. This allows researchers and policymakers to systematically study subjectivity. While this analysis explores different viewpoints in a specific population, it does not provide insights on the distribution of these viewpoints in that population, since the sampling is based on making sure that every viewpoint is found, instead of using a representative sample to elicit the viewpoints.

Furthermore, the preference elicitation methods that use rating methods are based on people's own judgement. Respondents in these studies are for instance asked about the strength of their preferences on a certain scale. These methods are suitable to be used in contexts where the kinds of preferences that people may have are already known, rather than in contexts where more explorative preference research is needed.

Indifference methods aim to change the value of an attribute of one of the choices until a point is found in which the respondent is indifferent between multiple choices. An example is a contingent valuation analysis. In a contingent valuation analysis, people are asked to directly report their willingness to pay to buy a certain good or service, or willingness to accept to give up a certain good or service. The advantage of this method is that it is relatively simple to conduct, and that it can be adapted to a wide range of contexts. A disadvantage is that there is evidence of disparity between WTP and WTA (see Rotteveel, Lambooij, Zuithoff, van Exel, Moons, de Wit [7] for evidence of the WTP/WTA disparity in a healthcare context). This suggests that the method is unable to capture true value, since various indifference methods lead to different estimates. Furthermore, the method is prone to various biases.[8] Respondents could for instance give strategic answers. They can respond with extreme values (for instance zero, or very high values) if they have very strong opinions of a good or if they do not agree with the

question itself (protest answers), or they can respond with a lower WTP if they think it affects the price that they will need to pay themselves. In the case of public provisions – when the provision of a good may be dependent on the responses but the price will not – they may in contrast overstate their WTP to increase the chance that a public good will be provided. In the context of measuring preferences on the allocation of public budgets, another critique is that the individuals' choice may not reflect how they want public policies to change.[9] Their individual contribution to the public good may also be perceived as negligible, and thus their choices may not reflect what we would like to measure.

In order to overcome the critique on measuring preferences on public budget allocation by contingent valuation analysis, collective WTP was designed. Nyborg [10] shows that in this method, subjects are asked what they believe everyone should pay to ensure a certain government project rather than what they would like to pay themselves. This resolves the coordination problem associated with private WTP. However, private money and public money are not always directly substitutable; people may for instance believe that a project should be funded by existing government funds rather than by extra tax income collected from the public. Willingness to allocate public budget (WTAPB) experiments were designed to let participants express that government funds and their own money have different purposes. In a WTAPB experiment, the subjects are asked to allocate the governmental budget to public goods, given that the government decides to allocate a certain amount of public budget. Their incomes are thus not impacted by their choices. Mouter, Van Cranenburgh, Van Wee [11] also empirically show that participants express different preferences in a WTAPB setting compared to a collective WTP setting. This suggests that participants feel that government funds and private money have different purposes.

In addition to WTAPB, the Participatory Value Evaluation (PVE)[9] method was designed to let participants have the option to advise the government against allocation (some) budget to

the projects shown and shift the remaining budget to next year. In a Fixed budget PVE, individuals are asked to allocate a certain public budget over a set of policy choices. It is thus also possible to shift a part of the budget to next year. In a Flexible budget PVE, individuals are also free to increase or decrease this public budget by increasing taxes or by shifting part of the budget to next year, which combines the WTAPB with the collective WTP approach. Social welfare effects are measured through the elicitation of individuals' preferences over the allocation of the public budget and their private income.[9] Until now, PVE experiments have mainly been used in transportation contexts.

Research context: disinvestment in The Netherlands

The research context of this study is decision-making on active disinvestment of healthcare in The Netherlands. To be able to manage the rising healthcare costs, various measures could be considered. More budget could be created to be able to afford the rising healthcare costs, or the rise of healthcare costs could be diminished. One of the possible austerity measures is cutting the budget for standard reimbursements of healthcare. This can be done by designing a stricter policy for new treatments to be reimbursed. Another option that has been suggested, but has not been commonly implemented is disinvestment of healthcare interventions; (partially) stopping reimbursement of certain healthcare interventions that are currently reimbursed because of, among others, ineffectiveness or insufficient value for money.[12,13]

Decision-making on the reimbursement of healthcare interventions is, however, internationally seen as a difficult process.[14] Even in the presence of very strong evidence that a certain medical treatment is insufficiently (cost-)effective, withdrawing a certain healthcare intervention from the reimbursed healthcare system is challenging. Politicians may feel that withdrawal may lead to societal discontent, and medical staff may feel that it is a patient's right to receive a treatment when it is available. The importance of careful considerations and

legitimate decision-making is therefore high. This increases the notion that the decisions about healthcare reimbursement should be guided by patient welfare. The accountability for reasonableness framework by Daniels [17] shows that a fair process is essential to social acceptance of the decisions made. This includes "transparency about the grounds for decisions; appeals to rationales that all can accept as relevant to meeting health needs fairly; and procedures for revising decisions in light of challenges to them."[17]

To prevent societal protest it is important to take into account the societal perspective when evaluating reimbursement decisions. However, there is no single perspective that can be taken into account when considering the public opinion. Instead, people differ amongst each other in what they consider to be important criteria for reimbursement decisions.[15,16] Considering that we are living in relatively pluralist societies now, it is no surprise that there is reasonable disagreement about the priorities of healthcare.[17]

There are a few studies which have focused on people's preferences regarding reimbursement or active disinvestment of healthcare. Specific research on the public opinion regarding cancer drugs in Canada shows that Canadian citizens highly value effectiveness, treatments that restore patients' independence, mental health and general well-being. This study also finds that people find it important the decision-making processes and its results are transparent.[18] This research, however, focuses specifically on cancer drugs. In a more general context, public views on principles for healthcare priority setting in ten European countries have been studied by van Exel, Baker, Mason, Donaldson, Brouwer, EuroVa [15]. Using a Q methodology, they find that five distinct viewpoints were identified on the priorities of healthcare. Given that no single equity principle was found, they stress the importance of the process of decision making, in which the plurality of viewpoints should be reflected. Currently, there is still a lack of information regarding what people generally consider to be important

aspects in disinvestment decisions. It is thus challenging to take into account the 'acceptable rationales'

In this analysis, we focus on active disinvestment in The Netherlands. In The Netherlands, all residents are entitled to a basic healthcare package. Everyone is obliged to have a basic health insurance, for which they pay a premium every month. Citizens are free to choose their own health insurer, which compete with each other on price and quality. The basic healthcare package is similar for each citizen. People with a low income receive an allowance to pay for the premium. Next to the premiums, everyone has a minimum own risk of 385 Euros. People can choose to increase their own risk with steps of 100 Euros until a maximum of 885 Euros. If people choose to increase their own risk, they receive a discount on their healthcare premium.

The Dutch Minister of Health, Welfare and Sports determines what healthcare is reimbursed by the basic care package and what is not.[19] The Ministry of Health, Welfare and Sports and the National Healthcare Institute[19] advise the Minister on this. The National Healthcare Institute gives advise based on four criteria: effectiveness, cost-effectiveness, necessity and feasibility.[20] The Ministry then advises the Minister taking into account this assessment, and other considerations that are deemed important, such as political and societal support.

Expensive innovative drugs are regularly first placed in the 'sluice'. These products can only be insured via the basic healthcare package when there are sufficient guarantees for responsible use and good evidence of efficacy, when reasonable financial arrangements can be made between the supplier and the Minister of Health, Welfare and Sports, and the National Healthcare Institute has advised on their inclusion in the basic package. This 'sluice' is designed to reduce the costs of expensive drugs as part of healthcare expenditure.

While a critical assessment of the inclusion criteria of healthcare into the basic healthcare packaged has been discussed widely, active disinvestment of healthcare interventions is less common. In the Netherlands, recent attempts of active disinvestment of healthcare have differed in their outcome. Rotteveel, Lambooij, Zuithoff, van Exel, Moons, de Wit [21] show with qualitative research that support for disinvestment among stakeholders strongly affected the outcome of the disinvestment process, while no evidence was found for a consistent role of formal considerations (such as effectiveness, cost-effectiveness, necessity and feasibility) in the disinvestment process.

Public support thus has played an important role in the outcome of past cases of disinvestment attempts. However, knowledge on the criteria that are deemed important by the public in the disinvestment process are rare. Therefore, two studies on this have been conducted in the last year; one Q-methodology study which elicits views on active disinvestment in The Netherlands[4], and one PVE which explores the relative importance of treatment characteristics in the context of active disinvestment in The Netherlands. Both studies try to provide insights on what the Dutch public deems important reimbursement criteria in the context of active disinvestment. In this study, we explored to what extent the findings of those two analyses are related.

Methodology

The findings of an earlier conducted Q methodology will be linked with findings of a Participatory Value Evaluation (PVE).

Q-Methodogy

The Q-methodology study by Rotteveel, Reckers-Droog, Lambooij, de Wit, van Exel [4] was conducted in 2019. The goal of the study was "to identify and describe views of citizens in the

Netherlands on the considerations they find relevant in the context of active disinvestment of healthcare interventions from the basic benefits package"[4]. The detailed methodology can be found in Rotteveel, Reckers-Droog, Lambooij, de Wit, van Exel [4].

The Q-methodology study found the following four viewpoints: 1) right to necessary care, irrespective of the costs, 2) right to necessary care, but within reasonable limits, 3) deliberate and fair spending of the healthcare budget, and 4) transparent and consistent decision-making. The descriptions of these viewpoints can be found in Table 1.

Table 1: Viewpoints elicited by Rotteveel, Reckers-Droog, Lambooij, de Wit, van Exel [4]

Viewpoint	**Description[1]**
Viewpoint 1: right to necessary care, irrespective of the costs	Everyone has the right to receive necessary care, especially if one has a serious illness. Costs are not important in this decision. Treatment with a small health gain must also be reimbursed. Medical doctors know best which treatments are necessary. Therefore, they must determine which treatments should be reimbursed.
Viewpoint 2: right to necessary care, but within reasonable limits	Everyone has the right to receive necessary care. However, in order to be able to pay for all necessary care, choices need to be made. It is therefore important that the effectiveness of healthcare is objectively established. Furthermore, people have to pay for their own healthcare if they are financially able to do this. Reimbursement of healthcare is especially important if the treatment prevents care in the future, or if there is no alternative treatment available. Stopping the reimbursement of healthcare is only acceptable if there is enough societal support.
Viewpoint 3: deliberate and fair spending of the healthcare budget	We have to be deliberate with our healthcare expenditure. Therefore, we should only reimburse healthcare that is truly necessary and effective. Furthermore, treatments should not be too expensive in relation to the health gain. Preventing diseases is also important. In addition, patients should not decide on reimbursement. The doctor or the hospital should determine whether healthcare is really necessary. It should not matter whether the healthcare provider has already earned back its investment.

[1] These descriptions are translations of the description shown to respondents in the Participatory Value Evaluation.

Viewpoint 4: transparent and consistent decision-making	It must always be clear why certain care is or is not reimbursement. This makes citizens understand the decision that was made. Furthermore, decisions on reimbursement should always be made in the same way. This ensures that everyone is treated equally. Whether someone can financially contribute to his/her treatment should not play a role. The most important thing is that effective care remains reimbursed. We must ensure that healthcare remains affordable.

Results from a Q-methodology study cannot be generalised to a larger population, since the sampling method in a Q-methodology study is designed with the purpose of representativeness of viewpoints in a population, rather than representativeness of a population itself. Furthermore, only a small sample is used. In order to find out the distribution of these viewpoints in a population, this should be tested separately in a sample representative of the respective population.

Representativeness of viewpoints

In this study we asked respondents to what extent they agree with each of the four viewpoints found by Rotteveel, Reckers-Droog, Lambooij, de Wit, van Exel [4], which allows us to quantitatively show the distribution of these viewpoints in the Dutch population. The last questions of the questionnaire containing the PVE task elicit the extent to which the respondents agree with the four viewpoints. The questionnaire explained that respondents were going to be shown four descriptions of viewpoints regarding active disinvestment of healthcare treatments. For each of these descriptions, they should indicate to what extent on a 7-point Likert scale they agree with that statement. The viewpoints were shown in a random order. Since the questionnaire took a long time to fill in, these Q-methodology questions were optional, as to prevent respondents from dropping out from the questionnaire.

In order to explore the distribution of the four viewpoints in the Dutch population, we matched the respondents to different classes based on their response patterns on the four Q-methodology viewpoints.

Participatory Value Evaluation

The attributes and its levels in the PVE were determined based on literature review and expert knowledge. The article by Rotteveel et al. on the PVE results will further elaborate on these choices. The attributes and its levels of the PVE are shown in Table 2.

Table 2: Attributes and levels of PVE including restrictions

Attribute	**Levels**
Number of patients treated	• 200 • 400 • 1,000 • 15,000 • 90,000 • 400,000
Costs per patient per year	• €100.00 • €200.00 • €1,000.00 • €5,000.00 • €50,000.00 • €90,000.00
Budget impact per year	Number of patients treated * Costs per patient per year, limited to resulting levels between 1 and 90 million Euros. The sum of the budget impact per year of all treatments together should be at least 190 million Euros.
Quality of life before treatment	• 35% • 55% • 75%
Quality of life after treatment	• 40% • 55% • 80% • 85%

Quality of life improvement	Quality of life after treatment – quality of life before treatment, restricted to values between 0 and 45 percent.
Remaining life expectancy before treatment	• 1 • 5 • 15 • 25
Remaining life expectancy after treatment	• 1.5 • 6 • 15 • 30
Improvement in remaining life expectancy	Remaining life expectancy after treatment – remaining life expectancy before treatment, restricted to values between 0 and 15 years.
Availability of alternative treatment	• An alternative treatment is available that treats the disease • An alternative treatment is available that treats disease symptoms but not the disease itself • This is the only available treatment for the disease and disease symptoms
Average age of the patient group	• 10 • 35 • 55 • 75, only combined with a life expectancy before treatment of 1, 5 and 15 and a life expectancy after treatment of 1.5, 6 and 15.

The PVE was conducted online in April and May 2020, collecting a sample of about 1,200 observations who completed the PVE task. The respondents were recruited by panel organisations CG research and Kantar Public, using probability sampling to get a sample representative of the Dutch population. In this fixed budget PVE, respondents were shown eight medical treatments, differing in attribute levels, including budget-impact. They were asked to choose what treatments they would cut if they would need to cut at least 100 million Euros of the public healthcare budget. They were allowed to save more than 100 million Euros. If respondents chose to save more than 100 million Euros, they were told to assume that in the

future, fewer budget cuts would be necessary. The treatments were hypothetical and unlabelled. There were sixty different designs of portfolios of treatments that were shown to respondents. Before and after the PVE task, several background questions were answered by the respondents. The PVE results were analysed using the statistical software R. The PVE was analysed using a portfolio model.[2] The assumption of this model is that respondents choose the portfolio from which they gain the highest societal utility.[9] In our analysis, the portfolio PVE model assumes that every respondent chooses to keep the portfolio with the combination of treatments which gives them the highest possible societal utility, compared to any other possible portfolio of treatments that respect the budget constraint. In order to more intuitively interpret the results of the PVE we use kept treatments (i.e. treatments that were not selected to be disinvested) as the dependent variable.

The PVE analysis results in specific parameters for each of the attribute values. With these parameters it is also possible to derive the relative preferences for the attribute values by comparing the sizes of the estimates of the attributes.

The utility[3] that an individual *n* perceives from portfolio *p* is given by the formula:

$$U_{np} = \sum_{j=1}^{J} y_{nj}\, U_{nj} + \alpha_B \left(\sum_{j}^{J} y_{n,-j} \cdot c_j - S_0 \right) + \varepsilon_{np},$$

Where:

- $y_{n,j}$ is a variable that is 1 if treatment *j* is kept (so not removed from the basic healthcare package), and 0 otherwise. Similarly, $y_{n,-j}$ is a variable that is 1 if treatment *j* is removed from the basic healthcare package, and 0 otherwise (if kept).
- U_{nj} is the utility derived from treatment *j*.

[2] The R script for this analysis was written by José Ignacio Hernández (TU Delft).
[3] The formulas in this paper were provided by José Ignacio Hernández (TU Delft).

- α_B is the marginal utility of savings additional to the required 100 million Euros. This is calculated by the total cost of discarded treatments minus the minimum savings ($c_j - S_0$).
- ε_{np} is the stochastic error term.

The utility of each possible hypothetical treatment can then be derived by the formula:

$$U_{nj} = \sum_{k=1}^{K} \beta_{jk} x_{njk} + \eta_{nj},$$

Where:

- β_{jk} is the marginal effect of attribute *k* on the utility of treatment *j*.
- x_{njk} is the portfolio of the kept treatments with treatments *j* and attributes *k*.
- η_{nj} is the stochastic error term.

We assume that the stochastic term is zero. The model assumes that an individual chooses a portfolio if the utility that they derive from this particular portfolio is greater than any other possible portfolio that they could choose taking into account the budget constraint. The probability that individual *n* chooses portfolio *p* over all possible portfolios *q* is thus given by:

$$P_{np} = P\big(U_{np} > U_{nq}\big), \forall\, q \in Q,$$

Where *Q* is the set that contains all possible portfolios of policies that satisfy the minimum amount of savings.

If we assume that the stochastic error term ε_{np} follows an extreme value distribution, P_{np} is similar to the multinomial logit probability where all portfolios contained in the variable *Q* are part of an individual's choice set. The multinomial logit estimate can be interpreted as: a unit increase in the attribute value increases the utility of keeping this treatment in the basic healthcare package by its respective parameter estimate.

Association between Q-Methodology and PVE

A large part of the respondents who finished the PVE task also filled in the questions on the Q-methodology perspective. This allowed us to explore whether respondents with different viewpoints showed different choice patterns in the choice experiment. Hypotheses were made based on the viewpoints and factor arrays described in Rotteveel, Reckers-Droog, Lambooij, de Wit, van Exel [4]. Only the parts of the descriptions that could be connected with one of the attributes were used for the hypotheses. The hypotheses can be found in Appendix 1. Since the exploration of such preference methods has not been done before, we have explored two different approaches to explore the relationship between Q-Methodology and PVE: a latent class analysis and a subgroup analysis of the PVE experiment.

Latent class analysis

A latent class analysis is a way to quantitatively analyse response patterns. A latent class cluster analysis (LCCA) identifies clusters of choice patterns, and maximizes homogeneity within the clusters and heterogeneity among the clusters. The difference with traditional clustering approaches is that LCCA is a model-based technique based on similarity in response patterns, rather than just based on the distance between respondents.[22] The model is specifically fit for the population from which the sample is collected. It can also be applied to variables of different scale types, and therefore can include both qualitative as well as quantitative attributes and levels. The analysis yields groups of individuals that make a similar trade-off among attributes, or focus on similar attributes. The LCCA model then probabilistically assigns respondents to a cluster. [23]

LCCA is a cluster analysis that has been used in the context of a PVE before (see for example Volberda [24]). In these earlier applications the software Latent Gold was used. Since we did not have access to this software, we planned to do the LCCA in R using the R package

PoLCA. The attributes and its levels were included as indicators to the model. The optimal number of clusters could then be determined using the Bayesian Information Criteria (BIC) and the bivariate residuals (BVR). The lower the BIC, the better, and the higher the BVR, the better. In order to explore the relationship between these response patterns and the extent to which respondents agree with the presented viewpoints, we planned to include the viewpoints as covariates to the model after we established the optimal number of clusters. Wald statistics and p-values would be used to determine whether the indicators (attributes) and covariates (Q-methodology responses) were significant, as well as differed significantly across clusters.

Subgroup analysis

Since it was uncertain whether the LCCA would be feasible, we also performed a subgroup analysis of the PVE results, in which the subgroups were based on the Q-methodology responses. To create the subgroups, we matched the respondents to each viewpoint that they agreed the most with, as long as they at least somewhat agreed with the statement (factor ≥ 5). This means that respondents who, for example, agree to the same extent with viewpoint 2 and 3, but do not agree with viewpoint 1 and 4, were matched to both viewpoint 2 and 3. Alternatively, a respondent who does not agree with any of the viewpoints was placed in a separate subgroup.

We then conducted the portfolio analysis for each subgroup, and compared the parameter estimates. Since a large part of the respondents were matched to more than one viewpoint, it was not possible to properly compare the estimates. Therefore, we also estimated the parameter estimates for a group of respondents who can be uniquely matched to only one viewpoint. This allowed us to better compare the choice patterns of these groups, despite the smaller sample sizes.

Results

In total, 3387 respondents opened the survey, and 1224 (36%) respondents completed it. A total of 1002 respondents filled in at least one of the questions on the Q-methodology viewpoints. We are unsure whether respondents who responded to only part of the Q-methodology questions simply do not agree with the viewpoints they did not respond to, or responded only partly because they were tired. Since we do not have data available on the time respondents took to answer these questions, and the order in which they saw the questions, we exclude these respondents from the analysis. 714 of the respondents responded to each question in this section and are included in the main analysis. This sample contained 21.1% of the respondents who opened the survey, and 58.3% of the respondents who completed the survey.

Descriptive statistics

Table 3 shows the descriptive statistics of the sample of 714 respondents. Compared to the population statistics provided by the Central Bureau of Statistics Netherlands (CBS) and the literature, we can see that the sample is relatively young and relatively highly educated compared to the general Dutch population. The income distribution is relatively similar to the income distribution of the Dutch population.[25] Based on the other sociodemographic statistics and the averages of education, our sample is relatively representative of the Dutch population.

Table 3: Descriptive statistics of sample with complete answers on the PVE task and all the Q-methodology questions

Sociodemographic statistics			
	Sample		Dutch Population*
Gender	Number	Relative	
Men	319	44,7%	49,7%
Women	392	54,9%	50,3%
Else	3	0,4%	0,0%
Age			

18 - 25	148	20,7%	12,5%
26 - 35	174	24,4%	15,8%
36 - 45	109	15,3%	16,3%
46 - 55	116	16,2%	18,0%
56 - 65	123	17,2%	16,7%
66 - 75	39	5,5%	13,7%
> 75	5	0,7%	10,1%
Education			
Basisschool	7	1,0%	10,0%
Praktijkonderwijs, LBO, VMBO-K/B, MBO 1	59	8,3%	12,0%
MAVO, VMBO-T/G, Havo-/VWO-onderbouw	79	11,1%	8,6%
Havo, VWO, Atheneum, Gymnasium, MBO 2/3/4	269	37,7%	37,1%
HBO-/WO-bachelor	182	25,5%	19,6%
HBO-/WO-master, doctoraal	115	16,1%	11,1%
No answer	3	0,4%	1,5%
Income			
< 1000	56	7,8%	
1000 - 2000	171	23,9%	
2000 - 3000	177	24,8%	
3000 - 4000	101	14,1%	
4000 - 5000	73	10,2%	
> 5000	41	5,7%	
No answer	95	13,3%	
Health			
Mean	76.6	87	
Standard deviation	18.2	17	
n = 713 (1 NA)			

*Population statistics are based on StatLine [26] (gender and age), StatLine [27] (education) and Versteegh, Vermeulen, Evers, De Wit, Prenger, Stolk [28] (health).

Representativeness of viewpoints

Since a Q-methodology study does not provide evidence on the distribution of viewpoints in a population, this study also contributes by giving a quantitative view of the distribution of viewpoints regarding active disinvestment of healthcare in the Dutch general population. Table

4 shows the distribution of responses on a 7-point Likert scale for each viewpoint. We show response statistics for the group of individuals who have reported the extent to which they agree with a viewpoint for at least one viewpoint. Appendix 3 shows a frequency table of all response patterns to the Q-methodology questions.

Table 4: Distribution of responses on the Q-methodology questions

	Responses to the Q-methodology statements (n = 1002)							
	No answer	Strongly disagree	Disagree	Somewhat disagree	Neutral	Somewhat agree	Agree	Strongly agree
Viewpoint 1: Right to necessary care, irrespective of the costs	92 (9.2%)	50 (5.0%)	120 (12.0%)	132 (13.2%)	124 (12.4%)	217 (21.7%)	201 (20.0%)	66 (6.5%)
Viewpoint 2: Right to necessary care, but within reasonable limits	118 (11.8%)	53 (5.3%)	93 (9.3%)	85 (8.5%)	160 (16.0%)	233 (23.3%)	215 (21.4%)	45 (4.4%)
Viewpoint 3: Deliberate and fair spending of the healthcare budget	104 (10.4%)	64 (6.4%)	109 (10.9%)	112 (11.2%)	157 (15.7%)	204 (20.3%)	201 (20.0%)	51 (5.1%)
Viewpoint 4: Transparent and consistent decision-making	103 (10.3%)	29 (2.9%)	47 (4.7%)	71 (7.1%)	98 (9.8%)	187 (18.7%)	323 (32.2%)	144 (14.3%)

The responses show that, relatively, a lot of the respondents strongly agree with viewpoint 4. Compared to the responses for the other viewpoints, there are relatively a lot of respondents who (somewhat) disagree with viewpoint 1. For the other two viewpoints the distribution of responses is quite similar.

Table 5 shows the correlation between the viewpoints. All correlations are positive and statistically significant at, at least, the 10% significance level. This tells us that respondents who agree to a high (low) extent with one viewpoint are also more likely to agree to a high (low) extent with the other viewpoints. Despite the correlations being statistically significant, they

are relatively low (corresponding to the rule of thumb for interpreting correlation sizes in Hinkle, Wiersma, Jurs [29]).

Table 5: Correlation and significance levels between responses on the Q-methodology questions*

	Viewpoint 1	Viewpoint 2	Viewpoint 3	Viewpoint 4
Viewpoint 1		0.101*	0.188***	0.184***
Viewpoint 2	0.101*		0.246***	0.143***
Viewpoint 3	0.188***	0.246***		0.185***
Viewpoint 4	0.184***	0.143***	0.185***	

* Significance: * corresponds to 10%, ** to 5% and *** to 1% significance level.

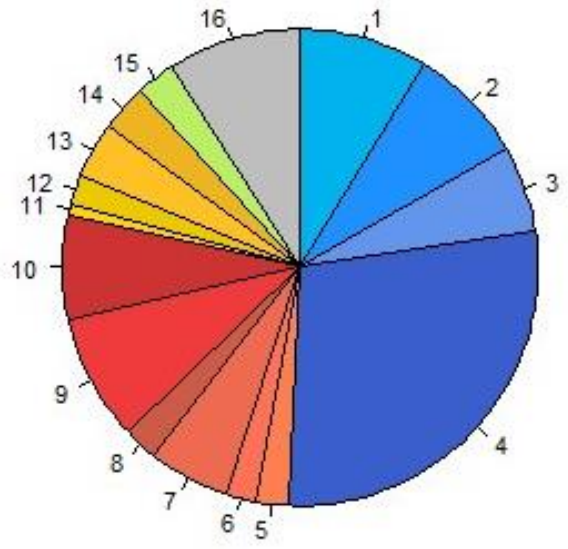
By assigning respondents to different classes we further explored the responses to the Q-methodology viewpoints. There is no way to perfectly assign respondents to a specific perspective as respondents may agree with multiple viewpoints, and it is unclear whether differences in the extent to which people agree with the viewpoints are reflected well enough in the 7-point Likert scale. Therefore, we used two different methods of matching people to the four viewpoints.

In this analysis we matched respondents to the viewpoint with which they agree the most. Other conditions were that respondents should at least 'somewhat agree' with the viewpoint (factor ≥ 5). If there are multiple viewpoints with which they agree the most, they were assigned to a mixed class of those viewpoints. There were only 65 respondents who, at maximum, were neutral (factor ≤ 4) about all four viewpoints. These respondents were thus not assigned to any viewpoint, and are in class 16 (no viewpoint).

Table 6: Matching of respondents to Q-methodology viewpoints*

Class	Viewpoints	Number of respondents	Share of total respondents
1	1	62	8.7%
2	2	57	8.0%
3	3	42	5.9%
4	4	202	28.3%
5	1 2	15	2.1%
6	1 3	15	2.1%
7	1 4	39	5.5%
8	2 3	17	2.4%
9	2 4	61	8.5%
10	3 4	51	7.1%
11	1 2 3	5	0.7%
12	1 2 4	15	2.1%
13	1 3 4	26	3.6%
14	2 3 4	23	3.2%
15	1 2 3 4	19	2.7%
16	None	65	9.1%
Total		714	100.0%

*Figure 2: Pie chart of share of respondents per matching class**



*Note: Respondents are assigned to the viewpoint with which they agree the most. Respondents are only assigned to viewpoint if they at least somewhat agree with the viewpoint (factor ≥ 5). If there are multiple viewpoints they are assigned to a mixed class of those viewpoints. If they do not agree with any viewpoint, the respondent is assigned to class 16 (none).

Table 6 shows the matching of the respondents. Figure 2 shows the share of respondents per class in a pie chart. It shows that about half of the respondents (50.9%) prefer a single viewpoint over the other three viewpoints. 27.7% of the sample agrees most to the same extent with two viewpoints, and 9.6% agree most to the same extent with three viewpoints. 2.7% of the sample agree to the same extent with all four viewpoints, and score at least a factor of 5 on all of those four viewpoints. 65 respondents (9.1%) are at maximum neutral (factor ≤ 4) about all four viewpoints. There are only 22 (3.1%) respondents who, at maximum, somewhat disagree (factor ≤ 3) with all four viewpoints. The matching classes with two or three viewpoints show that there is a high number of respondents who agree with viewpoint 4, compared to the other viewpoints. The second matching method, which assigns respondents to a viewpoint if they at least somewhat agree with a viewpoint, can be found in Appendix 2.

Participatory Value Evaluation

The Participatory Value Evaluation was analysed using the portfolio model. The results will be described and discussed in detail by Rotteveel et al.

34 respondents did not satisfy the budget constraint of 100 Million euros, but were classified as respondents who completed the experiment. These respondents were excluded from the analysis. Table 7 shows the results of this choice experiment for the other 680 observations. The variable 'marginal utility of additional savings' was fixed to be able to interpret the parameter estimates of the other attributes.

Table 7: Estimation results of Participatory Value Evaluation using the Portfolio model (dependent variable: kept treatments)*

	Portfolio Model
Age of the patient	-0.008***
	(0.001)
Additional Life Expectancy	0.064***
	(0.005)

Additional Quality of Life	0.019***
	(0.002)
There is an alternative treatment available that treats the symptoms of the disease**	-0.175**
	(0.065)
There is an alternative treatment available that treats the disease**	-0.390***
	(0.064)
Observations	680
Log-likelihood	-3333.921
AIC	6679.8423
BIC	6706.9749

* Significance: * corresponds to 10%, ** to 5% and *** to 1% significance level. Standard deviation is in parentheses.
** Reference category: There is no alternative treatment available.

The significance levels of the attributes show that every attribute significantly contributed to respondent's utility functions. The signs of the attribute estimates all correspond to what the project group expected prior to the analysis based on literature and experience. The older the patient group, the less likely that a treatment is kept. The larger the improvement in life expectancy and quality of life, the more likely the treatment is kept. If there is an alternative treatment available, respondents are less likely to keep the treatment. If that alternative treatment does not only treat the symptoms of the disease, but also the disease itself, respondents are even less likely to keep the treatment.

The sizes of the estimates describe the utility that respondents derive from every specific parameter if they keep the treatment. If the age of the patient group increases with one year, the utility that respondents derive from keeping this treatment decreases with 0.008. Thus, if the patient group age increases with twenty years, the utility that respondents derive from this treatment decreases with 0.064. This corresponds to the utility increase that people derive from a treatment that increases the life expectancy with one additional year. These two changes in treatment attribute thus cancel each other out.

Other interesting observations are that respondents derive three times as much utility from a life expectancy increase of one year compared to an increase of 1% in quality of life (0.064 compared to 0.019). In addition, if an alternative treatment is available that treats the disease itself, respondents derive twice as much utility from discarding this treatment compared to if this alternative treatment would only treat the symptoms of the disease (-0.390 compared to -0.175).

Relationship between Q-methodology and PVE

The association between the responses in the PVE task and the Q-methodology questions were studied by conducting a latent class analysis and doing a subgroup analysis.

Latent class analysis

The latent class analysis of the PVE results of this particular choice experiment turned out to be unfeasible at the moment. The possibility to analyse choice patterns in a PVE by conducting LCCA is under development, but has not yet been applied to an unlabelled, multiple-design PVE. Considering the complexity of the design of this particular PVE, we decided to focus on the subgroup analysis to explore the relationship between the responses in the PVE and the Q-methodology questions.

Subgroup analysis

In order to perform the subgroup analysis we first assigned respondents to different classes. The classes we matched the respondents to in Table 6 are relatively small. In order to be better able to interpret the results, we therefore assign respondents to each viewpoint that they are matched to in Table 6. A respondent who is assigned to class 6 (viewpoint 1 and 3) is for example assigned to both group 1 and group 3. The advantage of this is that the sample sizes are large

enough to meaningfully interpret the PVE estimates. The disadvantage is that, since a large part of the respondents belong to multiple classes, we cannot properly or statistically compare the PVE estimates of these subgroups.

The subgroups that we study are displayed in Table 8. We see that more than half of the respondents who filled in all the Q-methodology questions agree the most with (amongst others) viewpoint 4. The share of total respondents who agree the most with (amongst others) viewpoint 1, viewpoint 2 or viewpoint 3 is similar. Table 9 shows the estimates of the PVE portfolio model for each of the five subgroups.

Table 8: Subgroups

	Number of respondents	Share of total respondents
Viewpoint 1	196	27.5%
Viewpoint 2	212	29.7%
Viewpoint 3	198	27.7%
Viewpoint 4	436	61.1%
No viewpoint	65	9.1%

The minimal difference in the results show that respondents with different viewpoints show similar response patterns. This suggests that the preferences that respondents show in the choice experiment are separate from the general viewpoints that they agree with. We do observe small differences between respondents who agree with viewpoint 4 and respondents who agree with the other viewpoints. Respondents who agree most with viewpoint 4 gain less utility from the effectiveness (improvement in life expectancy and improvement in quality of life) of the treatment compared to the other respondents, and gain more utility if an alternative treatment is available compared to the other respondents. Yet, these differences are small.

Table 9: Subgroup analysis of Participatory Value Evaluation using the Portfolio model (dependent variable: kept treatments)*

	Viewpoint 1	**Viewpoint 2**	**Viewpoint 3**	**Viewpoint 4**	**No viewpoint**
Age of the patient	-0.008***	-0.007***	-0.007***	-0.009***	-0.004
	(0.002)	(0.002)	(0.002)	(0.001)	(0.004)
Additional Life Expectancy	0.069***	0.055***	0.055***	0.075***	-0.003
	(0.010)	(0.009)	(0.009)	(0.007)	(0.016)
Additional Quality of Life	0.018***	0.019***	0.019***	0.024***	0.001
	(0.004)	(0.004)	(0.004)	(0.003)	(0.006)
There is an alternative treatment available that treats the symptoms of the disease***	-0.253*	-0.112	-0.112	-0.163*	-0.341
	(0.125)	(0.119)	(0.119)	(0.083)	(0.212)
There is an alternative treatment available that treats the disease***	-0.269*	-0.346**	-0.346**	-0.289***	-0.674**
	(0.125)	(0.116)	(0.116)	(0.082)	(0.215)
Observations**	186	199	187	425	61
Log-likelihood	-907.4446	-984.1279	-984.1279	-2042.9917	-297.1353
AIC	1826.8893	1980.2557	1980.2557	4097.9834	606.2705
BIC	1846.2438	2000.0155	1999.6424	4122.2959	618.9358

* Significance: * corresponds to 10%, ** to 5% and *** to 1% significance level. Standard deviation is in parentheses.

** The number of observations is lower than the total number of respondents in the subgroup. This is due to the exclusion of respondents not adhering to the budget constraint of cutting the budget with at least 100 million Euros.

*** Reference category: There is no alternative treatment available.

Part of the reason that we find minimal differences may be the fact that 40% of the respondents belong to more than one category. Differences in response patterns are thus by definition limited, since we are comparing partly similar groups of people. Therefore, we repeat this subgroup analysis for the groups of respondents who agree most with only one viewpoint (class 1 to 4 in Table 6) or no viewpoint (class 16 in Table 6). The comparison of the PVE estimates between those classes are shown in Table 10.

Table 10: Subgroup analysis of Participatory Value Evaluation using the Portfolio model (dependent variable: kept treatments) for respondents who uniquely agree with viewpoint 1 or viewpoint 4.*

	Viewpoint 1	**Viewpoint 2**	**Viewpoint 3**	**Viewpoint 4**	**No viewpoint**
Age of the patient	-0.004	-0.007	-0.019***	-0.010***	-0.004
	(0.004)	(0.004)	(0.004)	(0.002)	(0.004)
Additional Life Expectancy	0.080***	0.055**	0.082**	0.083***	-0.003
	(0.018)	(0.019)	(0.023)	(0.010)	(0.016)
Additional Quality of Life	0.001	0.008	0.019	0.025***	0.001
	(0.007)	(0.007)	(0.009)	0.004	(0.006)
There is an alternative treatment available that treats the symptoms of the disease***	-0.217	-0.073	-0.068	-0.229	-0.341
	(0.226)	(0.232)	(0.285)	(0.121)	(0.212)
There is an alternative treatment available that treats the disease***	-0.531*	-0.753**	-0.218	-0.402**	-0.674**
	(0.220)	(0.233)	(0.280)	(0.121)	(0.215)
Observations**	57	50	39	199	61
Log-likelihood	-279.2339	-246.9302	-181.8742	-984.6291	-297.1353
AIC	570.4678	505.8604	375.7485	1909.2581	606.2705
BIC	582.7261	517.3326	385.7298	1929.0179	618.9358

* Significance: * corresponds to 10%, ** to 5% and *** to 1% significance level. Standard deviation is in parentheses.
** The number of observations is lower than the total number of respondents in the subgroup. This is due to the exclusion of respondents not adhering to the budget constraint of cutting the budget with at least 100 million Euros.
*** Reference category: There is no alternative treatment available.

Table 10 shows that respondents who agree with viewpoint 1 get a positive utility from treatments that increase the remaining life expectancy, but their utility is not significantly affected by an improvement in quality of life. They do significantly value whether an alternative treatment is available, but only if this alternative treatment treats the disease itself. The utility of respondents who agree with viewpoint 2 increases significantly if a treatment is kept which increases life expectancy or if a treatment is omitted for which an alternative treatment is available that treats the disease. The utility of respondents who uniquely agree with viewpoint

3 decreases when they keep a treatment for older patients and increases when a treatment increases life expectancy. However, whether there is an alternative treatment available that treats the disease does not significantly affect the respondent's utility, while this is the case for respondents of all other subgroups. Respondents who agree with viewpoint 4 care significantly about the age of the patient (negatively), improvement in life expectancy and quality of life (positively) and whether there is an alternative treatment available that treats the disease (negatively). The utility of respondents who do not agree with any viewpoint is only significantly affected by treatments for which there is an alternative treatment available that treats the disease.

Since these groups are mutually exclusive, we can properly compare the PVE estimates between these two subgroups. We see that only the utilities of respondents who agree with viewpoint 3 and viewpoint 4 are significantly and negatively affected by keeping a treatment of for older patients, while we do not see this effect for the other subgroups. Respondents with viewpoint 3 lose more than twice the utility from keeping a treatment for older patients compared to respondents with viewpoint 4. All respondents who agree with one of the four viewpoints care significantly and positively about the life expectancy gains of a treatment. Participants who do not agree with any viewpoint, do not. None of the subgroups see a significant utility change if a treatment is kept for which an alternative treatment is available that treats the symptoms of the disease. Only when this alternative treatment treats the disease itself, the utility of the respondents in subgroups of viewpoint 1, 2, 4 and no viewpoint decreases when the treatment is kept. Especially the utility of respondents who agree with viewpoint 2 is strongly affected by this attribute.

Conclusions

Since there is only a very small group of respondents who do not agree with all four viewpoints we conclude that the viewpoints found by Rotteveel, Reckers-Droog, Lambooij, de Wit, van Exel [4] are a good and complete representation of the viewpoints on active disinvestment of healthcare in the Dutch general population. The relatively low correlations between the viewpoints also suggest that the viewpoints are distinguishable enough. This provides evidence that Q-methodology is a good method to infer viewpoints about a specific topic in a specific population. 40% of the respondents agree most to the same extent with more than one viewpoint. This suggests that the viewpoints found by Rotteveel, Reckers-Droog, Lambooij, de Wit, van Exel [4] are building blocks of people's opinion, rather than exclusive opinions that people can agree with.

The distribution of responses show that the majority of the respondents (61% of the respondents who answered every Q-methodology question) agree with the viewpoint that decision-making should be transparent and consistent. For each of the other viewpoints a group of about 20% of the respondents agree with the respective viewpoints. Taking into account the total size of the Dutch population this means that if we assume that the distribution of responses in this sample is representative of the distribution in the Dutch general population, each viewpoint is agreed to by a large group of citizens.

Methodologically, we conclude that conducting a latent class analysis based on the PVE results is under development, but not yet feasible for all PVE experiments. The subgroup analysis has allowed us to explore the association between the extent to which people agree with the four Q-methodology viewpoints and the preferences they show in the PVE. We find limited differences between the PVE estimates for the five subgroups. We find that respondents who agree with viewpoint 3 and viewpoint 4 prefer treatments for young patients, while the other subgroups do not. In addition, only the utility of respondents in subgroup of viewpoint 4

is significantly affected by the quality of life improvement of a treatment. Lastly, the utility of the subgroup of viewpoint 3 is the only utility that is not significantly affected if there is a treatment available that treats the disease itself. Since the estimates for all subgroups except for the subgroup of viewpoint 4 are based on a small sample, we cannot make strong conclusions based on these results.

Separately, some results of subgroup analyses do match the hypotheses (in Appendix 1). In almost all subgroups of the four Q-methodology perspectives in Table 10, the attributes that significantly affected the respective utility of the respondents were hypothesized to significantly affect their utility with the corresponding signs. However, there were also hypotheses that were not confirmed by the subgroup analysis. It was for instance hypothesized that the utility of respondents who agree with viewpoint 1 would increase if they would keep treatments which improve quality of life. We also hypothesized that respondents who agree with viewpoint 2 would care mostly about quality of life, but instead their utility was affected by life expectancy instead of quality of life. In the subgroup of viewpoint 3, it was hypothesized that whether an alternative treatment would be available that treats the disease would significantly impact their utility. However, this did not significantly affect the respondent's utility, while this is the case for respondents of all other subgroups. The utility of respondents who agree with viewpoint 4 is significantly affected by more attributes than hypothesized; their utility for instance decreases if they keep treatments for older patients group, and if there is an alternative treatment available that treats the disease.

The differences in PVE estimates that we do find in the subgroup analysis thus do not correspond with the hypotheses, and do not make sense based on the descriptions. We therefore cannot claim that the hypotheses that we can accept are a result of having a certain viewpoint. Since we fixed the marginal utility of additional savings, we cannot make any claims regarding the relative importance of certain attributes compared to the budget impact that a treatment has.

The differences that we find in this subgroup analysis are thus limited and illogical. This can mean three things. Firstly, it can be the case that opinions and preferences in a choice experiment are largely unrelated. Secondly, it is possible that the differences that are found are drive by other factors (i.e. respondent characteristic) than differences in opinions. Thirdly, it can be that the measurement of preferences in both studies were not specific enough to distinguish differences between groups.

Discussion

The conclusions on the distribution of the Q-methodology viewpoints in the Dutch population imply that Q-methodology is a good method to explore what viewpoints are present on a specific topic in a specific population. The results tell us that the majority of people agree that decision-making regarding disinvestment should be transparent and consistent. When making complex disinvestment decisions, it is thus important to avoid that decisions are made behind closed doors. If they are, decisionmakers should report what decision they made and why, so citizens understand why this specific decision was made. Furthermore, people who agree with this viewpoint think that decisions on reimbursement should always be made in the same way. This calls for the use a univocal set of criteria, so that decisions are always made equally and transparently. Another specific opinion of respondents who agree with viewpoint 4 is that whether an individual can financially contribute to his/her treatment should not play a role. This implies that the government should not use different decision criteria for relatively low-cost healthcare than for more expensive healthcare. An opinion that is shared by all viewpoints and that is also reflected by the choices in the PVE is that effective healthcare should always be (or remain) reimbursed. This should thus be remain the primary criterium in decision-making on disinvestment if the government would like to incorporate the public's opinion.

In addition, we find that the differences in PVE estimates for the various subgroups are not plausibly related to the viewpoint descriptions. Therefore, we conclude that we do not find evidence that the findings of the Q-methodology study and the PVE findings are related. For policymaking this means that the methods should not be used to replace the other. It is important to explore the relationship between different methods of preference measurement so we know what kind of method derives what results. Since we do not find a plausible relationship between the Q-methodology and the PVE findings, this suggests that they should be used for different reasons. Q-methodology is a suitable method to explore viewpoints in a population when there is no extensive knowledge yet on what the people find important. However, when there is already a clear view on what people care about, for example based on earlier research or direct communication from the public, a PVE can be used to explore people's preferences in more detail. The analysis of a PVE results in the relative importance of characteristics, and thus more clearly provides people's preferences in comparison with other factors they find important. This more clearly provides guidelines for policymaking than the general viewpoints found in the Q-methodology. Without a clear idea on the factors that people find important, it is however challenging to properly design a PVE which includes the right attributes and levels. Depending on the existing knowledge regarding preferences, different methods can thus provide different useful insights. In this experiment, we used a PVE with hypothetical, unlabelled treatments so that the findings can be generalized to general decision-making on active disinvestment of healthcare. However, in the case of a specific budget allocation, labelled PVEs can be used to explore citizen's preferences in more detail, so that the findings can be easily applied in the policymaking process.

The strengths of our study are that firstly that we were able to explore the link between a Q-methodology study and a PVE using a sample of respondents who have both finished the choice task, as well as reported the extent to which they agree with the Q-methodology

viewpoints. We could therefore make use of microdata without needing to merge data sources. Furthermore, the relatively large and representative sample of Dutch population has allowed us to better explore the distribution of the viewpoints found in the Q-methodology study by Rotteveel, Reckers-Droog, Lambooij, de Wit, van Exel [4].

Next to the strengths, the study has various limitations that should be addressed. Firstly, it could be the case that there was selective drop-out. It is plausible that especially people who agree with viewpoint 1 (everyone has the right to necessary care, irrespective of the costs) would not fill in the survey because they do not agree that the healthcare budget should be cut. This would underestimate the share of people who agree with viewpoint 1, and could lead to bias in the relationship between the Q-methodology study and the PVE. Since the subgroup analysis does not a show that respondents who agree with viewpoint 1 have a logical different choice patterns than the respondents who agree with the other viewpoints (especially viewpoint 2 and 3), we do not expect that this selective drop-out would have led to a large bias in the analysis. Furthermore, there were also respondents who were excluded from analysis even though they completed the PVE task, since the questions following the PVE task were optional. If the characteristics of these respondents are different than the characteristics of the respondents who did fill in the Q-methodology questions, this has probably led to bias in the findings about the distribution of the Q-methodology viewpoints in the Dutch population. To further explore this bias, the characteristics of the respondents who dropped out can be compared to the characteristics of respondents who did complete the full survey at a later stage.

Secondly, the sample is relatively young and highly educated compared to the Dutch population. This may have biased our results. Further analysis on the characteristics of the different subgroups and matching classes is necessary to explore the potential bias in our results.

Thirdly, the subgroup analysis in Table 10 did not allow for proper comparison between response patterns. For that, a different kind of analysis or a bigger sample would be necessary.

A bigger sample would allow us to explore the sixteen possible classes in Table 6 separately, without having the issue of low sample sizes. However, for this a very large sample would be necessary, and this would most likely not be the most efficient way to further explore the association between Q-methodology and PVE findings. Other ways to further research this are explored below. In addition, the subgroup analysis in Table 10 can also be compared statistically, for instance by using an ANOVA-test. The ANOVA test compares estimates between the subgroups for every attribute. Considering the complexity of the PVE results and time limitations, this test was not conducted in this study. Since the ANOVA would include many inferences, a multiple testing problem may however occur, leading to false positives. Another option would be to do pairwise comparisons between subgroups using t-tests. This would, however, not provide us with a complete comparison of the subgroups.

Fourth, the fact that we did not find convincing evidence that respondents who agree with different viewpoints have different choice patterns in the PVE could also be the result of inconsistency between the design of the two methods. The choices that respondents could make in the PVE did not reflect all the differences between the four viewpoints. The description of most viewpoints, for instance, include whether it is up to the doctor, the government or the public to decide whether healthcare should be reimbursement. In addition, the attributes and levels of the unlabelled hypothetical treatments do not provide any information on whether the intervention treats preventable or communicable diseases, whether the healthcare provider has already earned back their investment, or whether there is any societal support for disinvesting a treatment. The Q-methodology study shows that these are aspects that some people care about in decisions regarding active disinvestment of healthcare. It could be the case that we did not find a convincing relationship between the two methods to measure preferences, since the respondents could not express important aspects of their opinion in the PVE task.

Another disadvantage of the current analysis of the PVE is that because the respondents were obliged to save at least 100 million Euros, the relative importance of budget impact per year is hard to measure. It is possible to measure the marginal utility of additional savings, but in a portfolio analysis of a PVE one attribute needs to be fixed in order to be able to interpret the other parameter estimates. In this analysis, we chose to fix this attribute of additional savings, which does not allow us to compare the importance of attributes like effectiveness of the treatment with the importance of additional savings. Further analyses could fix another attribute level while estimating the model to be able to make this comparison. Furthermore, some hypotheses were too specific, which meant that we were not able to test them using this analysis. Additional analyses that include the influence of life expectancy before treatment and quality of life before treatment, or the interaction between different attributes, would allow us to test these more complex hypotheses as well.

Lastly, the multiple discrete-continuous extreme value (MDCEV) model explained in Bhat [30] would be a more intuitive model to analyse the PVE model than the portfolio model used in this study, since the MDCEV model derives choice probabilities instead of utility estimates. However, considering the complexity of the MDCEV model this analysis was not conducted in this study. Further research can take this analysis into account, and potentially explore if the relationship between Q-methodology and PVE findings changes if a different model is used to analyse the PVE responses.

In order to tackle selective non-response or drop-out of respondents in surveys on sensitive topics like these we suggest that more attention is given to respondents who feel uncomfortable with the topic during the pilot tests. Those respondents may be able to pinpoint where they would quit the study. In collaboration with communication experts, the text can then be improved to avoid drop out and to increase awareness of the necessity of filling in the survey. In addition, to avoid drop-out that occurs later in the survey, close attention should be paid on

the length of the questionnaire. Long questionnaires can be mentally draining to complete. Therefore, in some studies it may be better to split the questionnaire into several separate surveys if it is acceptable that the distribution of both surveys results in separate samples. If the goal of this study would have been to solely explore the representation of the viewpoints found by the Q-methodology study in the Dutch population, it would have thus been better to explore this via a separate, shorter survey. Furthermore, we strongly encourage to start questionnaires like these with questions on sociodemographic and other background characteristics, so that if drop-out occurs throughout the survey it is possible to analyse which people dropped out of the choice task. The characteristics of the respondents who dropped out in this study will be analysed by Rotteveel et al. in their paper on the general PVE results.

This study and its results also call for further research. During the time in which we conducted this study we came across various other ways to further explore the relationship between Q-methodology and PVE responses. Firstly we hope that a latent class analysis as described in the methodology section of this study will be feasible to conduct in the coming years. The complexity of PVE designs like the one used in this analysis did not allow us to apply a latent class analysis at the moment. However, the application of a latent class analysis has been shown to work in other PVE experiments so far, and we expect that the current pace of development of PVE analyses will allow a latent class analysis to be possible in the coming years, considering that the use of PVEs in order to involve citizens in policymaking is gaining in popularity. Secondly, another promising method to explore the relation between choice patterns in a PVE and other respondent characteristic is the use of interaction terms in the PVE analysis. To fully analyse the effect of viewpoint responses on PVE choices, an interaction would be added for each possible pairwise combination of attributes and Q-methodology viewpoints. The hypotheses in Appendix 1 could also be used to decide on relevant interactions to include, to limit the number of interactions added to the model. Both of these proposed

methods would allow to explore the full variation in Q-methodology responses, rather than a matching-based method used in the current study. Furthermore, a benefit of these analyses is that we would explore the relationship between Q-methodology and PVE using the full sample instead of splitting the sample in subgroups. This approach can also be used to explore whether the differences found in Table 10 are driven by other respondent characteristics.

Ideally, further studies that try to answer a similar research question to ours ('To what extent are opinions and preferences in a choice experiment associated?') would make use of a PVE that is specifically designed based on the findings of a Q-methodology study. This would allow the exploration of all the differences between viewpoints. The Q-methodology that is analysed in this paper elicits viewpoints on a rather complex topic, and the study finds that there are many factors that citizens find important. It is unfeasible to include all those factors in a single PVE study. In order to match the findings of a Q-methodology study with a PVE study, both studies would probably require to have a very specific topic, since a PVE study is only effective with a limited amount of attributes and levels.

We recommend further analysis of the way that different methods to measure preferences of citizens are related. In order to properly explore and compare preferences in different populations and on different topics, it is essential to know what is being measured and to what extent the findings can be compared.

References

1. Adang EMM. Investeringsfonds is niet de oplossing voor stijgende zorgkosten. *ESB.* 2020.

2. VTV. Zorguitgaven: Hoe ontwikkelen zich de zorguitgaven in de toekomst? https://www.vtv2018.nl/zorguitgaven. Published 2018. Accessed 2 July, 2020.

3. OECD. *Public funding of health care.* OECD;2020.

4. Rotteveel AH, Reckers-Droog VT, Lambooij MS, de Wit GA, van Exel NJA. Disinvestment of the reimbursement of healthcare interventions – A Q-methodology study on relevant considerations. Unpublished.

5. Soekhai V, Whichello C, Levitan B, et al. Methods for exploring and eliciting patient preferences in the medical product lifecycle: a literature review. *Drug Discov Today.* 2019;24(7):1324-1331.

6. Watts S, Stenner P. *Doing Q methodological research: Theory, method & interpretation.* Sage; 2012.

7. Rotteveel AH, Lambooij MS, Zuithoff NPA, van Exel J, Moons KGM, de Wit GA. Valuing Healthcare Goods and Services: A Systematic Review and Meta-Analysis on the WTA-WTP Disparity. *Pharmacoeconomics.* 2020.

8. Kouwenberg L. *Elicitation techniques in WTP/WTA research: Narrative literature review about the different techniques used in stated preferences research and their advantages and disadvantages.* 2018.

9. Dekker NMPKT. An introduction to Participatory Value Evaluation. *Tinbergen Institute Discussion Paper.* 2019.

10. Nyborg K. Homo economicus and homo politicus: interpretation and aggregation of environmental values. *Journal of Economic Behavior & Organization.* 2000;42(3):305-322.

11. Mouter N, Van Cranenburgh S, Van Wee B. Do individuals have different preferences as consumer and citizen? The trade-off between travel time and safety. *Transportation research part A: policy and practice.* 2017;106:333-349.

12. Harris C, Allen K, King R, Ramsey W, Kelly C, Thiagarajan M. Sustainability in Health care by Allocating Resources Effectively (SHARE) 2: identifying opportunities for disinvestment in a local healthcare setting. *BMC Health Serv Res.* 2017;17(1):328.

13. Rooshenas L, Owen-Smith AL, Hollingworth W, Badrinath P, Beynon C, Donovan JL. "I won't call it rationing…". *An ethnographic study of healthcare disinvestment in theory and practice.* 2015;128:273-281.

14. MacKean G, Noseworthy T, Elshaug AG, et al. Health technology reassessment: the art of the possible. *Int J Technol Assess Health Care.* 2013;29(4):418-423.

15. van Exel J, Baker R, Mason H, Donaldson C, Brouwer W, EuroVa QT. Public views on principles for health care priority setting: findings of a European cross-country study using Q methodology. *Social Science & Medicine.* 2015;126:128-137.

16. Reckers-Droog V, van Exel J, Brouwer W. Who should receive treatment? *PLoS ONE.* 2018;13(6).

17. Daniels N. Accountability for reasonableness, Establishing a fair process for priority setting is easier than agreeing on principles. *BMJ.* 2000;321:1300-1301.

18. Bentley C, Peacock S, Abelson J, et al. Addressing the affordability of cancer drugs: using deliberative public engagement to inform health policy. *Health research policy and systems.* 2019;17(1):17.

19. Nederland Z. Over ons. https://www.zorginstituutnederland.nl/over-ons. Published 2020. Accessed 2 July, 2020.

20. Zwaap J. *Package advice in practice: delibrations for arriving at a fair package.* Zorginstituut Nederland; 5 September 2017 2017.

21. Rotteveel AH, Lambooij MS, Zuithoff NPA, van Exel NPA, Moons KGM, de Wit GA. What influences the outcome of active disinvestment processes in healthcare? A qualitative interview study on five recent cases of active disinvestment. Unpublished.

22. Kroesen M. Latent class (cluster) models. In: Faculty of Technology Policy and Management, Delft University of Technology; 2019.

23. Vermunt JK, Magidson J. Latent class cluster analysis. *Applied latent class analysis.* 2002;11:89-106.

24. Volberda LJ. *Analyzing citizen's views on new spatial-infrastructure projects: from the average view towards various clusters within the Participatory Value Evaluation Method*: Transport Infrastructure and Logistics, Delft University of Technologie; 2019.

25. CBS. Inkomensverdeling. CBS. https://www.cbs.nl/nl-nl/visualisaties/inkomensverdeling. Published 2019. Updated 13 November 2019. Accessed 2 July, 2020.

26. StatLine. Bevolking; geslacht, leeftijd en burgerlijke staat, 1 januari. StatLine. https://opendata.cbs.nl/statline/#/CBS/nl/dataset/7461BEV/table?fromstatweb. Published 2019. Updated 2 December 2019. Accessed 2 July, 2020.

27. StatLine. Bevolking; hoogstbehaald onderwijsniveau en onderwijsrichting. StatLine. https://opendata.cbs.nl/#/CBS/nl/dataset/82816NED/table?dl=3948A. Published 2020. Updated 15 May 2020. Accessed 2 July, 2020.

28. Versteegh MM, Vermeulen KM, Evers SM, De Wit GA, Prenger R, Stolk EA. Dutch tariff for the five-level version of EQ-5D. *Value in health.* 2016;19(4):343-352.

29. Hinkle DE, Wiersma W, Jurs SG. *Applied statistics for the behavioral sciences.* Vol 663: Houghton Mifflin College Division; 2003.

30. Bhat CR. The multiple discrete-continuous extreme value (MDCEV) model: role of utility function parameters, identification considerations, and model extensions. *Transportation Research Part B: Methodological.* 2008;42(3):274-303.

Appendices

Appendix 1: Hypotheses

Viewpoint	Hypothesis	Hypothesis on relative importance
Viewpoint 1: right to necessary care, irrespective of the costs	Participants highly value a high quality of life improvement and a remaining life expectancy improvement. Participants choose to cut reimbursement of treatments with a high quality of life before treatment. Participants will not choose to cut reimbursement of treatments for which there are no alternative treatments available. Participants do not take budget impact per year into account. Participants will not choose to cut reimbursement of treatments for young patients with a low remaining life expectancy before treatment.	Participants who strongly agree with this perspective are highly likely to have a response pattern in which quality of life improvement, remaining life expectancy improvement and alternative treatment are highly valued rather than budget impact per year.
Viewpoint 2: right to necessary care, but within reasonable limits	Participants will not choose to cut reimbursement of treatments of which there are no alternative treatments available. Participants do not choose to cut reimbursement of treatments that have a high quality of life improvement, especially for people with a low quality of life before treatment. Participants do not take the budget impact per year and the costs per patient per year into account.	Participants who strongly agree with this perspective are highly likely to have a response pattern in which alternative treatment, quality of life before treatment and quality of life improvement are highly valued rather than budget impact per year and costs per patient per year.

Viewpoint 3: deliberate and fair spending of the healthcare budget	Participants choose to cut reimbursements of treatments with high costs per patients per year and a high budget impact per year, especially of which an alternative treatment is available that treats the disease. Participants choose to cut reimbursements of treatments with a high quality of life before treatment. Participants will not choose to cut reimbursement of treatments for young patients with a low remaining life expectancy before treatment.	Participants who strongly agree with this perspective are highly likely to have response patterns in which costs per patients per year, budget impact per year, alternative treatment (that treats the disease) and remaining life expectancy before treatment are highly valued. Treatments for young patients with a low remaining life expectancy before treatment are also highly valued.
Viewpoint 4: transparent and consistent decision-making	Participants choose to cut reimbursements of treatments with a low quality of life and life expectancy improvement. Participants do not necessarily take quality of life before treatment into account. Participants do not take the costs per patient per year and the budget impact per year into account. Participants do not choose to cut reimbursement of treatments that have a high quality of life improvement, especially for people with a low quality of life before treatment.	Participants who strongly agree with this perspective are likely to have response patterns in which quality of life improvement and life expectancy improvement are highly valued rather than quality of life before treatment, costs per patient per year and budget impact per year.

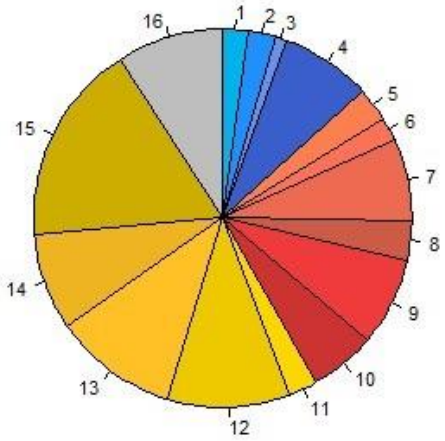
Appendix 2: Alternative matching of respondents to Q-methodology classes

Next to the matching method used in the main analysis, we have also matched respondents to every viewpoint that they at least slightly agree to (factor ≥ 5). The distribution of respondents is shown below.

Table: Matching of respondents to Q-methodology viewpoints*

Class	Viewpoints	Number of respondents	Share of total respondents
1	1	15	2.1%
2	2	17	2.4%
3	3	7	1.0%
4	4	56	7.8%
5	1 2	20	2.8%
6	1 3	15	2.1%
7	1 4	50	7.0%
8	2 3	24	3.4%
9	2 4	54	7.6%
10	3 4	40	5.6%
11	1 2 3	18	2.5%
12	1 2 4	75	10.5%
13	1 3 4	76	10.6%
14	2 3 4	59	8.3%
15	1 2 3 4	123	17.2%
16	None	65	9.1%
Total		714	100.0%

*Figure: Pie chart of share of respondents per matching class**



*Note: Respondents are assigned to every viewpoint with which they at least somewhat agree (factor ≥ 5). If there are multiple viewpoints they are assigned to a mixed class of those viewpoints. If they do not agree with any viewpoint, the respondent is assigned to class 16 (none).

Appendix 3: Frequency table of all response patterns to the Q-methodology questions

Frequency table of responses to Q-perspectives											
Index	**Freq**	**Index**	**Freq**	**Index**	**Freq**	**Index**	**Freq**	**Index**	**Freq**	**Index**	**Freq**
1111	4	2516	1	3566	6	4674	1	5663	2	6624	1
1112	1	2533	1	3567	1	4767	1	5664	2	6626	2
1115	1	2541	1	3573	2	5116	1	5665	2	6627	2
1136	1	2545	2	3574	1	5126	1	5666	4	6632	1
1146	1	2546	2	3623	1	5155	1	5667	4	6633	1
1147	1	2555	1	3636	4	5156	2	5677	3	6634	1
1167	1	2556	3	3641	1	5166	2	5726	1	6635	1
1213	1	2557	1	3644	1	5177	1	5755	1	6636	2
1216	1	2564	1	3646	1	5216	1	5756	1	6637	2
1263	1	2565	2	3654	1	5225	1	5763	1	6642	1
1276	1	2566	2	3655	2	5226	3	5765	2	6645	1
1313	1	2567	2	3656	1	5232	1	5766	1	6646	2
1316	1	2613	1	3657	1	5235	2	5767	1	6647	1
1326	1	2616	1	3662	1	5236	1	5777	1	6655	2
1334	1	2622	1	3663	1	5253	1	6116	1	6656	2
1412	1	2625	1	3665	1	5255	1	6123	1	6657	1
1416	1	2626	1	3666	4	5256	1	6157	1	6663	2
1426	1	2627	1	3667	1	5266	2	6177	1	6664	1
1455	1	2636	1	3676	1	5275	1	6211	1	6665	1
1465	1	2644	1	3713	1	5326	2	6224	1	6666	8
1467	1	2645	3	3717	1	5334	1	6236	1	6667	2
1526	1	2646	3	3743	1	5355	1	6237	1	6675	1
1532	1	2651	1	3755	1	5356	2	6242	1	6676	1
1565	1	2655	1	3757	1	5362	1	6244	1	6677	3
1572	1	2656	2	4135	1	5363	1	6252	1	6726	1
1626	2	2662	1	4167	1	5375	1	6254	1	6735	1
1645	1	2664	1	4226	1	5414	1	6255	1	6745	1
1651	1	2666	2	4227	1	5415	1	6256	3	6747	1
1652	1	2667	2	4233	1	5416	1	6257	1	6767	1
1656	2	2754	1	4245	1	5426	3	6266	5	6776	1
1666	1	2756	1	4246	1	5435	3	6267	2	7111	1
1717	1	2763	1	4256	1	5436	3	6277	1	7127	1
1736	1	2766	1	4262	1	5444	1	6317	1	7144	1
1737	1	2767	1	4266	2	5446	1	6327	1	7152	1
1751	1	2777	1	4335	4	5447	1	6331	1	7156	2
2111	1	3115	1	4344	2	5455	1	6336	2	7157	1
2117	1	3123	1	4345	1	5456	2	6345	1	7216	1
2121	1	3214	1	4356	1	5463	1	6347	1	7225	1
2122	1	3226	2	4422	1	5464	1	6356	2	7256	1
2147	1	3227	2	4426	2	5466	2	6363	1	7277	1
2212	1	3235	3	4427	1	5467	2	6364	1	7326	1

2213	1	3236	1	4434	3	5475	1	6366	1	7345	1
2215	2	3255	1	4436	2	5517	1	6416	1	7377	1
2222	1	3321	1	4443	2	5523	2	6424	1	7415	1
2227	1	3326	1	4444	17	5524	1	6425	1	7416	1
2241	1	3327	1	4445	1	5526	2	6426	3	7417	1
2252	1	3332	1	4453	1	5527	1	6434	1	7422	1
2255	2	3333	3	4454	1	5533	1	6435	1	7466	1
2256	3	3336	1	4455	3	5535	3	6445	1	7474	1
2266	1	3344	1	4456	1	5536	6	6446	1	7522	1
2267	1	3355	1	4466	2	5544	1	6453	1	7536	1
2312	1	3356	3	4467	3	5545	2	6454	1	7537	1
2325	2	3427	1	4514	1	5546	4	6455	1	7542	1
2327	1	3434	1	4524	1	5552	1	6457	2	7544	1
2331	1	3435	1	4526	1	5555	6	6463	1	7546	2
2333	1	3443	1	4537	1	5556	3	6466	5	7555	1
2334	1	3444	2	4543	1	5557	1	6467	2	7556	1
2335	1	3446	1	4544	1	5565	3	6474	1	7575	1
2344	2	3456	1	4545	3	5566	9	6476	1	7577	1
2346	1	3462	1	4546	1	5567	1	6516	1	7622	1
2357	1	3517	1	4553	1	5576	1	6525	1	7647	1
2365	1	3525	1	4554	1	5613	1	6527	1	7652	1
2366	1	3526	1	4555	5	5614	1	6535	1	7663	1
2376	1	3534	1	4556	2	5615	1	6536	2	7664	1
2422	1	3535	3	4563	1	5616	1	6542	1	7665	1
2424	3	3536	1	4565	1	5617	1	6543	1	7666	3
2425	1	3537	1	4566	2	5625	1	6545	3	7667	1
2435	3	3545	1	4567	2	5626	1	6546	5	7675	1
2444	1	3546	4	4625	1	5636	2	6547	3	7676	2
2446	1	3547	1	4626	1	5637	1	6553	2	7677	2
2447	1	3553	1	4637	1	5642	1	6555	1	7737	1
2455	2	3555	4	4645	1	5643	1	6556	4	7747	1
2456	1	3556	6	4655	1	5646	6	6557	1	7766	1
2466	1	3557	2	4656	1	5654	2	6563	1	7777	5
2473	1	3562	1	4657	1	5656	11	6565	5		
2513	1	3564	1	4661	1	5657	7	6566	6		
2515	1	3565	1	4667	1	5662	1	6576	1		