USING A POINT SYSTEM IN THE MANAGEMENT OF WAITING LISTS: THE CASE OF CATARACTS*

Carmen Herrrero, José L. Pinto-Prades and
Eva Rodríguez-Mínguez**

WP-AD 2001-16

Correspondence to Carrmen Herrero, University of Alicante, Departamento de Fundamentos del Análisis Económico, Campus San Vicente del Raspeig, 03071 Alicante, e-mail: Carmen.Herrero@ua.es.

Editor: Instituto Valenciano de Investigaciones Económicas, S.A.
First Edition June 2001.
Depósito Legal: V-2647-2001

---

* The authors acknowledge financial support from Spanish Ministry of Health and Consumption and financial support from CICYT, PB97-0120.

** C. Herrero: University of Alicante and Ivie; J.L. Pinto-Prades: Pompeu Fabra University and University of Vigo; E. Rodríguez-Mínguez: University Pompeu Fabra.

USING A POINT SYSTEM IN THE MANAGEMENT OF WAITING LISTS: THE CASE OF CATARACTS

Carmen Herrrero, José L. Pinto-Prades and
Eva Rodríguez-Mínguez

ABSTRACT

The objective of this study is to explore the possibility of applying a point system as a guide to the management of waiting lists in National Health Systems. Following recent contributions in the axiomatic theory of justice, the ethical properties of a point system are illustrated. In addition, we present the results of an experiment whose objective was to develop a point system for cataract extraction, based on social preferences. The results of the experiment have shown that the analytic methods used here, of focus groups, interview-administered questionnaires, conjoint analysis, and rank-ordered logit, can be usefully combined to determine the total priority score for each patient.

KEYWORDS: Point System; Priority Criteria; Waiting Lists; Cataract; Conjoint Analysis; Rank-Ordered Logit Model.

1 Introduction

Health care is a commodity with special characteristics. It is understood that equal access to a basic package of health services is essential for a fair society. For this reason, in many countries health services are mainly provided by the State. Leaving the provision of health care entirely to market forces would lead to rather unfair situations. It would probably be deemed inhumane that some members of a society could not receive appropriate medical care because of their low incomes. National Health Systems (NHS) have therefore been developed to provide health care to all members of society, irrespective of their incomes. The right to receive health care is based on medical needs.

To ensure that no-one remains untreated, simply because of his/her low income, NHS are state funded and patients usually do not have to pay at the point of consumption. Achieving fairness, however, has its costs. In NHS, this cost is (often) reflected in the great demand for health services. Given that, for equity reasons, most western societies reject prices as a mechanism for solving the problem of resource allocation, other rationing mechanisms have to be used. There are several means for rationing demand, apart from prices. One is to exclude some potential patients from the system. Exclusion can be based on income, rejecting those who have a sufficiently high levels of income to pay for their own health care. Another option —which is not incompatible with the former one— is to exclude patients with certain types of illnesses, so that a minimum treatment package is defined and free medical care is only provided to those suffering from health problems that can be treated by interventions that are included in the package. The option most frequently used is not based on any kind of explicit exclusion, but on the implicit rationing mechanism of waiting lists.

The first two options have hardly been used. The first has been mooted in Italy, but it has never been put into practice up to the time of writting this paper. The second option, of providing a minimum package of treatments, was attempted in Oregon to determine funding priorities for Medicaid beneficiaries, but without much success. One of the criticisms of this system is that patients are heterogeneous and those suffering from severe health problems may be excluded.[1] Waiting lists are very much used in practice, but they also have several problems. One is that waiting times have been considered excessive for several medical procedures. The British government tried to address this problem in the Patient´s Charter, by establishing a maximum waiting time. However, the British Medical Association was critical of this provision. They argued that waiting time was not the key element in judging the management of waiting lists. Instead, they considered medical needs to be much more important. They suggested that *patients should be given severity scores when they were put on a waiting list, which would reflect their clinical priority and how quickly they should receive surgery.*[2]

An important step forward in the needs-based management of waiting lists were the reforms that were implemented in New Zealand. The government decided to address the rationing problem, not as in Oregon, i.e. by excluding some medical treatments from the basic package, but by choosing the patients on waiting lists with a higher level of need. Groups of experts elaborated a point system that was used to measure need in a broad sense. This system gave priority to patients on the basis of their clinical and social characteristics. For example, in the case of fertility services, the characteristics that were used to prioritize among patients were the probability of success, age, number of children, etc.

Basing the rationing of public services on a point system is not new, and many examples can be found in Elster (1992).[3] However, it has not been widely used in the management of health care waiting lists. The only clear experience has been the system used to allocate organs in the US (the UNOS system).[4], [5] The novelty of the New Zealand experience is that they used a point system for selective surgery over a wide variety of health problems.[6],[7]

In the belief that this approach has many positive features, we decided to study the possibility of developing and applying such a point system in Spain for a specific condition, namely cataracts, which was chosen simply for convenience. A further objective of this study was to examine some of the ethical aspects of the point systems as they are applied to health care waiting lists. This is done by providing, together with well-known properties, some novel results in the field of the axiomatic theory of justice. The aim of this secondary analysis was to examine whether the management of waiting lists, using a point system, could have a sound theoretical base. The final objective of the study was to show how social preferences can be incorporated into the estimation of points, so that the relative weight of each of the characteristics used to prioritize patients reflects (as far as possible) society´s opinion, i.e. the opinions of the tax-payers — and potential users— of the NHS. In this paper we show how this can be done using conjoint analysis, a methodology which is common in market research but has not been used in the management of waiting lists.

The paper is structured as follows. In Section 2, we provide with a theoretical foundation, from the point of view of the Theory of Justice, to the use of point systems to manage waiting lists. In Section 3, we describe the conjoint analysis methodology used to obtain a point system for cataracts based on social preferences. We conducted a survey among 100 people from the general population to elicit such preferences. We also used an innovative statistical method, namely the *rank ordered logit model*, to analyze the data. This kind of model has not been previously used in health care management, though it may be the most appropriate statistical model to use to determine the relative positions of patients in a waiting list. Section 4 shows the results of the survey and it is followed by a discussion on the study.

2 The point system. Theoretical foundations

Let us consider a particular class of allocation problems. The characteristics of these problems are:

- There is a single good, that comes in indivisible units (as, for example, positions in a college, tickets for a performance, grants, operations, etc.)
- There is a group of agents who each demand one unit of the good.
- Time is discrete. At any point in time, the total supply of the good is not enough to cover the total demand for it.
- The price of the good is fixed, and there are no possibilities for side payments to compensate the agents.

A particularly simple situation arises when all of the agents have, in principle, an equal right to be served. If this is the case, the principle of *first come, first served*, is normally applied, as in the case of tickets for a performance. Nonetheless, there are many situations in which not all of the agents have an equal right to be served (for example, in the college admission problem, agents with better qualifications are more likely to be admitted). This is also the case for queues arising in the case of health care supply.

When agents do not have all equal rights, we have to somehow select which agents should be served, according to their *right.* In so doing, it is frequent to select certain characteristics of the agent which are considered relevant in defining his right. For example, in the health care provision problem we may consider urgency, health deterioration, likelihood of recovery, etc. By combining all the possible levels of their characteristics, we define *types.* Thus, agents are classified into types and, in order to select who should be served, the agents' types are taken into account. Finally, it is important to design a procedure that is able to properly select the agents to be served. With these ideas in mind, let us now consider a more detailed account (technical definitions are showed in Appendix A).

There is a set $\mathcal{N}$ of potential agents, and at a point in time, a subset N of agents demand the good, and at that particular point in time there are a number of available units of the good, or, alternatively, a number of agents who can be served, s. A p*roblem* is a pair (N, s) such that $\#N \geq s$. Let $\mathcal{P}$ be the set of all possible problems. Our aim is to select, for any given problem, the set of agents to be served.

For every problem, an allocation criterion, Φ, selects the set of allocations (solutions) for such a problem. As a consequence, $\Phi(N, s)$ identifies at least one set of possible agents to be served.

2.1 Types. Priority. Priority relation

An obvious way of ordering and then selecting the agents in any particular problem is according to the order of their appearance. But taking only this into account is not a sensible way of serving the agents. Certain characteristics of the agent are generally considered relevant for allocation purposes. Such characteristics allow us to classify the agents into types in the following way. Let there be a finite set of relevant characteristics, T, labelled attributes, where each attribute $t \in T$ can have only a finite number of values, called levels. By combining all of the levels of the different attributes, the set of types, J, is obtained. So a type $j \in J$ is defined by a vector of levels, one for each attribute, $\chi_j = (x_{jt})_{t \in T}$, where x_{jt} indicates the level of attribute t in type j. Therefore, if the type of agent $a \in N$ is denoted as $\chi(a) = (x_{at})_{t \in T}$ —where x_{at} indicates the level of attribute t in the agent a— we shall say that an agent a is of type j if $(x_{at})_{t \in T} = (x_{jt})_{t \in T}$ for all t.

By classifying the agents into types, we avoid any other considerations that are not relevant to the assignment problem. If two agents, a and b, are such that $\chi(a) = \chi(b)$, we may interpret this as implying that a and b are indistinguishable *according to their relevant characteristics.* Imagine, for instance, that we face the problem $(\{a, b\}, 1)$, with $\chi(a) = \chi(b)$. Since a and b are indistinguishable, at first sight we cannot differentiate between allocating the only available unit of the good to a or to b. Both allocations would be equally fair. So, given that only the attributes are considered when deciding on allocation, an allocation criterion Φ that selects the set of *fair allocations* must discriminate amongst agents only if they are of different types. An allocation criterion Φ is said to be *anonymous* if it selects fair allocations according to the relevant data (type) only and not by the names of the agents.

Consider now a group of fair allocations selected by an allocation criterion. If an individual happens to belong to all of those allocations, he will certainly be served. If not, the relative number of fair allocations to which he belongs somehow measures his *relative right* to be served. Suppose that an individual leaves the problem. It should be fair that the relative right of all remaining individuals weakly increases. An allocation criterion is said to be *monotonic with respect to the population (pop-monotonic)* if whenever an individual disappears from a problem, everybody else's right weakly increases. [For a survey on this idea in several models, see Thomson (1995)[10]].

We have shown two basic principles of equity that any allocation criteria should satisfy. We now need to built-up a criterion which verifies these principles and allows us to solve the allocation problem (N, s).

Suppose that we have a weak preference relation ρ on the set of types J, so that for all χ_j and χ_k, $\chi_j \rho \chi_k$ if an agent of type χ_j has at least as much right to the good as an agent of type χ_k. Relation ρ, is normally known as a *priority relation.* Of course, ρ also induces a weak order on the set of agents in the obvious way: $a\rho b \iff \chi(a)\rho\chi(b)$. If $a\rho b$ and $b\rho a$ occur, we say that agent a and b *are in a par.*

Whenever we have a priority relation ρ, it can be used to built-up an allocation criterion, Φ^{ρ}, by simply selecting, for any problem (N, s), all allocations with exactly s individuals each, all of whom enjoy higher or equal priority to those who were left out of that allocation.

It can be shown that if an allocation criterion Φ is based on a priority relation ρ, then is *anonymous* and *pop-monotonic.* Furthermore, all anonymous and pop-monotonic allocation criteria come from a priority relation on the set of types. (Appendix A, Theorem 1). The previous axiomatic characterization of allocation criteria coming from priority relations is new, and can be viewed as an alternative to that presented in Young (1994) by using impartiality and consistency. Apart from the novelty of this characterization, we think that, in this context, anonymity and population monotonicity are easier to understand.

2.2 Point systems

A common way of ordering types, is by using a point system. A point system is an allocation criterion that induces a weak order ρ on J, in the following way: For every attribute $t \in T$, assigns weights to its levels, $v_t[\cdot]$ —where a higher weight indicates a more preferred level— and the priority of each type $\chi_j = (x_{jt})_{t \in T}$, is established as a function of the sum of the weights associated to his levels of the different attributes,

$$U(\chi_j) = \sum_{t \in T} v_t[x_{jt}]. \tag{1}$$

So, given two types, χ_j and χ_k, we say that type χ_j has at least as much right to the good as type χ_k if $U(\chi_j) \geq U(\chi_k)$. Conversely, any separable weak order on the set of types can be represented by a point system (Appendix A, Theorem 2).

For example, suppose that two attributes, *gender* and *age*, are considered to be relevant in ordering the agents in a queue. Our first attribute, $t = age$, takes on three levels: $x_{1t} = below$ 15; $x_{2t} = over$ 65; $x_{3t} = between$ 15 *and* 65. The second attribute, $t' = gender$, assumes two levels, $x_{1t'} = woman$, and $x_{2t'} = man$. Consider now that women should go before men, children (below 15 years old) should go before elderly people (over 65), and elderly people should go before people between 15 and 65. There are three customers: a, a 35-years-old woman; b, a 68 years old man, and c, a boy of 10 years. Consider that a and b are in a par, and should be served after c. Then a point system compatible with this preference relation would be the set of weights, $v_t(x_{1t}) = 3, v_t(x_{2t}) = 2, v_t(x_{3t}) = 1, v_{t'}(x_{1t'}) = 2, v_{t'}(x_{2t'}) = 1$, given that it verifies that $U(a) = U(b) = 3$ and $U(c) = 4$.

Additionally, because of the particular form of this utility function, it happens that ρ satisfies the property of separability. That is to say, there are no complementarities between the different attributes. To explain this idea, consider two types $\chi_1$and χ_2, that are identical in all but two attributes —say they differ in attribute 1 and 2. Consider two other types χ_3 and χ_4 that are like $\chi_1$and χ_2 in the first two attributes, and that are like each other —and different to $\chi_1$and χ_2— in all other attributes. The priority relation ρ is separable in attributes 1 and 2 if the priority between $\chi_1$and χ_2 is the same that between χ_3 and χ_4. For example, if $\chi_1 \rho \chi_2$ then $\chi_3 \rho \chi_4$. Relation ρ satisfies the property of separability if it is separable in all pair attributes.

Given that an allocation criterion can provide more than one possible allocation, an allocation rule compatible with Φ is nothing but a selection of Φ. To construct an allocation rule that is compatible with an allocation criterion we need a break-tie rule, to be able to choose one of the fair allocations prescribed by Φ. In the above mentioned example, a and b are in a par, and should be served after c. As such, faced with the problem $(\{a, b, c\}, 2)$, there are two possible *fair allocations*, namely, $\{c, a\}$ and $\{c, b\}$. For this problem, an allocation rule necessarily chooses a unique solution. We may choose any form of tie-breaker. For instance, by choosing any pre-established linear order on the agents in N: date of birth, name, or more in keeping with the idea of rights, in the case of waiting lists, their oder of arrival (first come, first served).

3 Method: conjoint analysis

The properties of point systems set out in the previous section justify their use as a tool for assigning an order to patients on a waiting list. To design a point system that could be used to manage waiting lists for cataracts, and which would take societal preferences into account, the technique of conjoint analysis (CA) was used.[12] Although CA has been widely used in market research since the mid-1970s, its use in the field of health economics has been minimal.[13],[14]

CA is used to elicit individuals' preferences for sets of multi-attribute alternatives (products, services, etc.). The technique is based on obtaining weights for the different levels of each attribute that are more consistent with an individuals's overall preferences for the set of alternatives. In this study, the alternatives will be patients who require a given health-care service, i.e., an operation for cataracts, and the attributes will be the characteristics of the paptient that are considered to be relevant in deciding their positions on a waiting-list, including severity, time on the waiting-list, etc. Beginning with overall preferences, our objective was to obtain a value for every level of each attribute, to be able to calculate a total priority score for each one.

In compliance with the CA methodology, the following steps were followed to obtain a point-system that could be used to prioritize patients for cataract-extraction: a) selection of attributes and levels; b) selection of combination of attributes (patient types) to be rated in the survey; c) selection of a method for data collection; d) questionnaire design; e) selection of participants; f) selection of the estimation method; and g) selection of tests to assess validity and reliability.

A) Selecting attributes and levels. On the basis of the existing literature [15], [16],[17] and after two 2-hour in-depth interviews with 4 opthalmologists, the following attributes and levels were selected for inclusion in the CA.

1. ***Visual incapacity (I).*** This attribute indicates the severity of visual impairment, and is measured by specialists using objective techniques to produce a rating of what is known as *visual acuity.* The concept of visual acuity. however, is very much a medical one, and its inclusion in a questionnaire designed for use in the general population would be impractical. To overcome this problem, and after consultating specialists involved in the study, the concept of visual acuity was converted into an attribute based on the number of activities that the patient could or could not perform. The greater the number of activities which the patient cannot perform due to their impaired vision, the greater their degree of impairment.

In order to develop this attribute, an index of functional impairment called the VF-14 [18] was used. The VF-14 measures patients' level of incapacity when performing 14 vision-dependent activities of daily living, such as driving, reading small print, or watching television. Using the VF-14 it was possible to rank each of the 14 activities included in the instrument according to the degree of visual capacity that was required to perform each activity.[19] This ranking permitted the design of an attribute —visual incapacity— which could be easily understood by members of the general population (see appendix). Four levels of visual incapacity were established: a) *mild*, b) *moderate*, c) *severe*, and d) *very severe.*

2. ***Limitations in daily activities (L).*** This attribute refers to the patient's capacity to perform the everyday activities that he or she performed before becoming visually impaired. The information provided by this attribute is different to what is provided by the visual incapacity attribute, as it is possible that the same level of visual incapacity mightl affect different individuals in very different ways, in terms of their everyday activities. For example, the effect of mild visual incapacity (unable to read small print or to drive) on a pensioner who hardly ever drove, and on a lorry-driver, would clearly be very different, with the latter being much more severely limited, in terms of everyday activities.

This attribute has 2 levels: a) *the patient has some problems in performing* *everyday activities (for example: working, doing housework, participating in family or leisure activities)* and b) *the patient is unable to perform the majority of his or her everyday activities.*

3. Likelihood of improvement (K). This attribute indicates the likelihood of the patient recovering full sight after the operation. For the sake of simplicity, only total or minimal recovery of sight are considered as possible outcomes. It should be remembered that cataracts do occur along with other eye diseases (macular degeneration, glaucoma, etc) which reduces the likelihood of a successful operation. The probability of success is also reduced by the presence of comorbid conditions, such as diabetes. This attribute therefore records the effectiveness of the operation, and according to the existing literature[20], [21] and expert opinion, the likelihood of improvement is divided into three levels: a) *moderate (50%)*, b) *high (75%)*, c) *very high (99%)*.

4. ***Patient's age (A).*** Given that this disease is prevalent among older patients, theis ages are categorized as follows: a) *50 years* b) *65 years* c) *75 years*, and d) *85 years of age.*

5. ***Time on waiting list (W).*** This attribute has been classified using 4 levels: a) *3 months*, b) *6 months*, c) *12 months*, and d) *18 months.*

B) Creating patient types for evaluation by respondents. The attributes and levels described above define 384 ($4 \times 2 \times 3 \times 4 \times 4$) patient types. Given the nature of the attributes chosen, the property of separability is assumed. This in turn allows the number of patient types to be evaluated to be reduced to 16, using an orthogonal fractional factorial design.[12],[22] This design ensures the absence of collinearity.

C) Choosing a data collection method. A variety of different preference elicitation methods exist. They include asking respondents to compare a series of two alternatives; to choose one alternative from a set of alternatives; to rate the full set of alternatives using one of a variety of scaling methods, or; ranking the set of alternatives. We consider the latter approach to be more suitable for the objectives of this study, because: first, it is important to know how all of the patients on the waiting list will be ordered; second, the use of a ranking method allows respondents to order the full set of alternatives, without an undue cognitive burden (the use of pair-wise comparisons to rank states would make the exercise unduly arduous); finally, given that the aim is to order patients on the waiting list, an ordinal measure will be sufficient.

D) Designing the questionnaire. The questionnaire consists of 4 different sections. The first part explains the study's objectives and the different attributes and levels included, and allows the respondent to ask questions about the study. In the second part, the respondent lists the 16 types of patient selected according to the order in which he or she thinks the patients should be operated on. To make the task easier, a description of each of the 16 patients is printed on a separate card, so that the respondent simply has to put them in order In the third part, the respondent is asked to put the 5 attributes —I, L, K, A and W— in order according to the importance they assign to each attribute in this prioritization exercise. The respondent is also asked whether there are any other patient characteristics they would consider relevant —as this information may well be valuable in the design of future surveys. Finally, the respondent is asked to review his order of the cards and to modify it if they consider it appropriate.

E) Selecting the respondents. The questionnaire was presented to 100 individuals drawn from the general population of Barcelona (47% men and 53% women) and representing different age groups (23% were between 18 and 29 years of age, 28% between 30 and 44, 31% between 45 and 64, and 18% were over 65). The questionnaires were administered during October 1999. The interview was arranged beforehand by telephone and was performed at the respondent's home by an interviewer who had been specially trained in the administration of the questionnaire.

F) Estimation method: rank–ordered logit model. Using the ranking provided by the participants as a starting point, we arrived at an evaluation (weight) for each level of each attribute, $v_t[\cdot]$, which is a necessary step towards obtaining a point system. Following expression (1) in section 2, we want to estimate the function $U\left(\chi_j\right) = \sum_{t=1}^{5} v_t\left(x_{jt}\right)$, for all j, denoted as U_j hereafter. This function provides the total priority score (total utility) assigned to a patient of type j.

To choose a more appropriate estimation model, we must take the type of information obtained into account: each interviewee person i, $(i = 1, 2, ..., 100)$, ranked 16 patients as those who should be operated on first (hereafter, the sub-index i refers to respondents interviewed). To analyze this type of information, Beggs[23] and Chapman[24] proposed using the rank–ordered logit model (ROL). This method not only takes the respondent´s most preferred alternative into account, as the ordered logit model does, but it also allows for the use of all the information contained in the ranking of alternatives. This additional information can provide more precise estimates of the unknown parameters.

The basic specification used to derive the ROL is the random utility model. This model considers that the preference relation of respondent i over the choice set J can be represented by a utility function U_{ij} that measures the non-observable utility assigned by respondent i to alternative j —defined as a vector of attributes χ_j. U_{ij} is assumed to be composed of two terms, one deterministic, $V\left(\chi_j\right)$, which represents the systematic component of the model, and the other stochastic, e_{ij}, which captures the measurement error. If one assumes that these two components are independent and additive, the random utility model may be written in the form

$$U_{ij} = V\left(\chi_j\right) + e_{ij} = V_j + e_{ij}. \tag{2}$$

We specify a particular linear form for V_j,

$$V_j = \chi_j \alpha = \sum_{t=1}^{5} x_{jt} \alpha_t, \tag{3}$$

where α is the parameters vector (or matrix) to be estimated. Eq.(3) is a good representation of the preference relation of the sample if, as we assume, the property of separability is verified.

The maximum likelihood method used to estimate the ROL allows us to obtain consistent and (asymptotically) efficient parameters. The computer program used to perform the estimations was LIMDEP 7.0.

Two models were estimated for V_j —Eq.(3)— in the first, being a *general model,* we obtained one parameter for each of the levels of the attributes. In this model, α_t is a vector that has as many elements as the t attribute has levels. As is usual in the literature, to avoid exact collinearity in the estimation, we parametrized the model so that one level of each attribute was excluded. Specifically, we excluded the levels labeled as "mild visual incapacity", "has some problems in performing everyday activities", "moderate likelihood of improvement", "85 years of age" and "3 months on waiting list". Therefore, the parameters associated with the remaining levels of an attribute must be interpreted in relation to the excluded level: a positive (negative) coefficient for a given level indicates a higher (lower) score for this level with respect to the excluded level.

>From estimated parameters in the general model, the relative importance (RI) of each attribute is obtained by dividing its range, i.e. the difference between the hiighest and lowest coefficients, between the sum of the ranges of all attributes.[26]

Secondly, we estimated the *adjusted model,* in which continuous variables are approximated by continuous functions. The advantage of this second model is its capacity to provide a valuation for those intermediate levels that were not considered in the survey. The consistent Akaike information criterion (CAIC) has been used to select the adjusted model. The CAIC allows us to compare non-nested models based on the log–likelihood function, penalising the decrease in degrees of freedom produced as the parameters increase. Its value is obtained from the expression $-2LnL\left(\alpha\right)+(1+Ln\ n)k$, where n is the number of observations and k is the number of parameters, the model with the smallest CAIC being the most preferred.

G) Validity and reliability.

1. Validity.

The validity of the results was tested in three ways:

Face validity. The signs of the estimated parameters were examined to determine whether they complied with *a priori* expectations.

Predictive validity. The correlation between the respondents' rank ordering and the model–generated order was tested at individual and aggregate levels. The Spearman rank correlation coefficient was used to calculate the correlation between each individual's rank ordering and the order generated by the model, and the average correlation coefficient was then calculated for all participants. To calculate the correlation at an aggregate level, individual orderings were aggregated using the Borda rule, and the Spearman rank correlation coefficient between this ordering and the ROL generated ordering was obtained.

Construct validity. The correlation between the rank ordering assigned directly to the 5 attributes by the respondents and that obtained from their relative importance was calculated. To do so, individual orderings of the attributes were aggregated using the Borda rule and the Spearman rank correlation coefficient between direct ordering and estimated ordering of the attributes was calculated. This comparison, although intuitively attractive, is in fact relatively unorthodox and the results obtained should be treated with caution.

2. Reliability. The reliability of the results was analyzed in two ways:

Internal consistency. Given the assumption of rationality in the prioritization process, respondents should give greater priority to greater incapacity, greater likelihood of improvement, greater limitations in performing usual activities and longer time on the waiting list. With regard to age, any preference is considered to be rational here. When comparing all pairs of alternatives that could be extracted from the 16 alternatives, 5 pairs were produced where one alternative was clearly superior to the other, i.e. age was the same in the two alternatives, and of the remaining attributes at least one was at a higher level and none at a lower level. In such a situation, the superior alternative should rationally receive a higher ranking. Internal consistency was tested by determining the percentage of respondents who were allocated rational rankings in each of these 5 cases.

Robustness to sample size. To determine the degree of robustness of the rankings to sample size, the Spearman rank correlation coefficient was obtained between the ranking generated by the model using the full sample and the ranking generated using half of the sample —randomly selected.

4 Results

Rankings of the 16 patients and the 5 attributes were obtained from all respondents and the interview lasted an average of 45 minutes.

4.1 Estimated models

The estimates generated by the general model are shown in table 1. All the coefficients estimated were significant, at the 1% level, and were in the expected direction. For example, as visual incapacity increases so does the value of the associated parameter, and hence the likelihood that the patient will be operated on earlier.

As can be seen, the two attributes with the greatest relative importance were visual incapacity (37%) and the patients age (28.7%), whilst the least important was the likelihood of improvement (6.6%).

Representing the estimated parameters graphically is of great help when proposing possible functional forms for attributes of a continuous nature. On the basis of the CAIC value, the adjusted model shown in table 2 was chosen. The adjusted model introduces a lineal adjustment for age, $v_a = \alpha_a A$, and a logarithmic adjustment for time on the waiting list (measured in months), and the likelihood of improvement , $v_w = \alpha_w Ln(W)$ and $v_k = \alpha_k Ln(K)$, respectively. The constant is not introduced because it is not significant. The coefficients of the qualitative variables are barely altered, and the continuous variables are in the expected direction. *Age* has a negative sign, and both *likelihood of improvement* and *time on the waiting list* have positive signs.

Based on the estimated parameters, the total priority score for any hypothetical patient can be obtained. For example, let two patients A and B, have the following attributes:

$$\chi_A = (very\ severe,\ has\ diffic.,\ 60\%,\ 60\ years,\ 4\ months),$$

$$\chi_B = (moderate,\ is\ unable,\ 80\%,\ 70\ years,\ 14\ months)\,.$$

Given that A's total priority score $\left(\widehat{V}\left(\chi_A\right) = 1,167\right)$ is greater than B's $\left(\widehat{V}\left(\chi_B\right) = 0,958\right)$, A will be given greater priority than B.

4.2 Validity and reliability

1. Validity.

Face validity. As mentioned, the parameters were in the expected direction.

Predictive validity. With regard to predictive validity, the Spearman rank correlation coefficient between ranking estimated using the ROL and the direct rankings of respondents was 0.67 at the individual level (average) —only 25% had a correlation below 0.64— and 0.98 at the aggregate level.

Construct validity. The ranking of the attributes obtained by aggregating the rankings provided by the respondents coincides with the ranking obtained by estimating relative importance using the ROL (correlation coefficient of 1). Although the comparison of both methods should be seen as merely a simple approximation, the agreement in the ranking reinforces the results obtained.

2. Reliability.

Internal consistency. Of the 5 pairs of alternatives where the consistency can be tested, one of the pairs was rationally ordered by 100% of the respondents, 3 of the pairs were rationally ordered by 99% of the respondents, and the remaining pair was rationally ordered by 94% of the respondents.

Robustness to sample size. The correlation coefficient between the ordering estimated by the model using the full sample and the other one using half of the sample was 0.96.

5 Discussion

There is currently considerable debate in the health economics literature regarding the characteristics that should be taken into account when deciding on a patient's place on a waiting list, with increasing agreement on the idea that more than merely medical characteristics should be taken into account. However, increasing the number of characteristics to be considered also makes the management of waiting lists more complicated. This paper has shown that the point- system may provide a useful tool for health care authorities when establishing priorities.

The study has also shown how CA can be used to estimate the relative weights (points) for attributes in patients awaiting surgery for cataract extraction. The means used to obtain such weights allowed for the incorporation of social preferences. Using the values assigned to different levels of the attributes it is possible to obtain the variations in score that are produced when the attribute levels are varied. For example, in Table 1 it can be seen that the increase in score produced by the move from *mild* to *very severe incapacity* cannot be exceeded by any change in level in any of the other attributes. For this to occur, changes on more than one attribute would be needed. For example, a change from age 85 to age 50 and a concurrent change from 3 months to 12 months on the waiting list.

In the case of cataracts, *visual incapacity* and *age* appear to be the most important attributes in the prioritization process. *Time on the waiting list*, on the other hand, is fourth in importance. This is a significant result, especially when it is borne in mind that, in Spain, *time on the waiting list* is the only variable considered to be of any relevance once the patient has been included on the list.

Although the results of the experiment are encouraging, in the sense that the tests of reliability and validity support the quality of the overall results, the study did have some limitations, which should be taken into account if the system were to be implemented in the health-care system in the future. On the one hand, patients of a similar type will have the same level of priority. Nevertheless, in some cases patients belonging to the same category may be perceived as being different, particularly if important attributes have not been included in the analysis. For example, the qualitative analysis revealed that whether the patient was responsible for others or not was important for some respondents. The same might be true if the number of levels per attribute were reduced to much. For example, given equality at the other levels, patients who cannot recognize people at close range might be considered to have a much higher priority than those who cannot read large print in a book. With respect to continuous variables, the points assigned to intermediate levels were obtained using an adjustment function. However, these values should be confirmed in future studies. Cards with intermediate levels, for example, could be incorporated into the survey to test the ability of the model to predict the relative position of those levels.

Similarly, another surprising result, which should be interpreted with caution, was the use of a logarithmic function to obtain the value assigned to *time on the waiting list.* This function reflects social preferences which assign increasingly lower values to each additional month that the patient spends on the waiting list. This acts as a brake on the weight that can be assigned to a patient according to the time he or she has spent on the waiting list and reduces the possibility that a long time on the waiting list might override the value of any other attribute. Nevertheless, it might also be the result of the respondents' having a *distorted* perception of the actual distance between two levels of a given attribute. For example, some respondents may perceive the difference between 6 and 12 months to be greater than the difference between 12 and 18 months. Which is why it would be useful to ask respondents directly whether they agree with assigning increasingly lower values to each additional month on the waiting list.

Another limitation of the study was the sample size. It should be remembered, however, that this study was a pilot study to determine the validity of the methodology employed to obtain a point-system. Although the test of robustness shows that the results were consistent within this particular sample,the sample size should be considerably increased to ensure that the results were representative of the society's preferences.

Finally, it would be useful, in future research, to specifically elicit preferences from all the groups directly implicated. This could be done in two ways. On the one hand, the focus groups could be carried with out not only health-care professionals as in the present study, but with patients, patients' relatives and members of the general population. This would help not only in the choice of attributes but also in the levels used with each attribute, which is a key feature in obtaining a points system. On the other hand, the survey should be administered not only to members of the general population, but also to health care professionals, patients and members of the patients' families. Although it might be assumed that the preferences of the different groups will be similar, this needs to be investigated in future studies.

In conclusion, this paper has shown that a point-system is an allocation criterion that verifies two basic principles of equity, anonymity and pop–monotonicity. This priority-setting method also makes the management of waiting lists transparent, given that the criteria used to prioritize patients are made explicit. It may help to resolve the dilemma faced by clinicians in trying to establish priorities among a very heterogeneous set of patients. The results of the experiment have shown that the analysis methods used here, of focus groups, interview-administered questionnaires, conjoint analysis, and rank–ordered logit, can be usefully combined to provide a means of obtaining a point system for the management of waiting lists for national health care systems, at least in the case of patients awaiting treatment for cataracts.

Appendix A

Let $\mathcal{N}$ be a population of potential agents. At a given point in time, a subset of agents, $N \subset \mathcal{N}$ demand one unit of a good each. Let s be the number of available units of the good. Then, a *problem* is a pair (N, s) such that $\#N \geq s$. Let $\mathcal{P}$ be the set of all possible problems. Our aim is to design a *rule* or *mechanism* such that for any given problem it selects the subset (or subsets) of agents to be served.

To be able to choose the set of agents who should be served, they are classified into types, according to a set of relevant characteristics (attributes). Consider a finite set of attributes, T, where each attribute $t \in T$ can have only a finite number of levels. A type j is defined by a vector of levels, one for each attribute, $\chi_j = (x_{jt})_{t \in T}$, where x_{jt} indicates the level of attribute t in type j. The type of the agent $a \in N$ is denoted as $\chi(a)$. Let J be the set of types.

Allocation criterion: is a correspondence, Φ, defined on $\mathcal{P}$ such that $\Phi(N, s)$ is a set of allocations, such that for all $A \in \Phi(N, s)$, (1) $A \subset N$, and (2) $\#A = s$.

Anonymous allocation criterion: Let (N, s), $(N', s) \in \mathcal{P}$ such that $\#N = \#N'$, and there exists a one-to-one function $\theta : N \to N'$, with $\chi(\theta(a)) = \chi(a)$, for all $a \in N$. Then, allocation criterion Φ is said to be *anonymous* iff for all $A \in \Phi(N, s)$, we have that $\theta(A) \in \Phi(N', s)$.

Relative right: For an allocation criterion Φ, for any problem (N, s) and for any agent $a \in N$, we can define the *relative right of agent a in* (N, s), *according to criterion* Φ, $\alpha^{\Phi}[a; (N, s)]$ as the number of allocations in $\Phi(N, s)$ containing a over the total number of allocations in $\Phi(N, s)$, that is,

$$\alpha^{\Phi}[a; (N, s)] = \frac{\#\{A \in \Phi(N, s) : a \in A\}}{\#\{A \in \Phi(N, s)\}}$$

Pop–monotonic allocation criterion: Let (N, s), $(N', s) \in \mathcal{P}$ such that $N' \subset N$. Then, an allocation criterion is said to be *monotonic with respect to the population* iff for all $b \in N'$, $\alpha^{\Phi}[b; (N', s)] \geq \alpha^{\Phi}[b; (N, s)]$.

Priority relation: Is a weak order ρ defined on the set of types.

A priority relation ρ also induces a weak order on the set of agents, in the obvious way. Whenever a priority relation ρ is defined on $\mathcal{N}$, an allocation criterion, Φ^{ρ}, can be constructed as follows: For all $(N, s) \in \mathcal{P}$, $\Phi^{\rho}(N, s) = \{A \subset N : \#A = s$, and $a\rho b$, for all $a \in A$, all $b \in N \backslash A\}$.

Theorem 1: *An allocation criterion* Φ *is anonymous and pop-monotonic iff there exists a priority relation* ρ *on the set of types,* J*, such that* $\Phi(N, s) = \Phi^{\rho}(N, s)$*, for all* $(N, s) \in \mathcal{P}$.

Proof: $\Longleftarrow$) By construction, Φ^{ρ} is anonymous: Let (N, s), $(N', s) \in \mathcal{P}$ such that $\#N = \#N'$, and there exists a one-to-one function $\theta : N \to N'$, with $\chi(\theta(a)) = \chi(a)$, for all $a \in N$. Let $A \in \Phi(N, s)$. Thus, $A \subset N$, $\#A = s$, and $a\rho b$, for all $a \in A$, all $b \in N \backslash A$. Consider now $A' = \theta(A)$. It happens that $A' \subset N'$, $\#A' = s$, and, for all $a' \in A'$, $a' = \theta(a)$, $a \in A$, and for all $b' \in N' \backslash A'$, $b' = \theta(b)$, $b \in N \backslash A$. Consequently, for all $a' \in A'$, $\chi(a') = \chi(a)$, and for all $b' \in N'\backslash A'$, $\chi(b') = \chi(b)$. Therefore, for all $a' \in A'$ and for all $b' \in N'\backslash A'$, we have that $a'\rho b'$, and thus, $A' \in \Phi(N', s)$.

Also, Φ^{ρ} is pop-monotonic: First, notice that ρ induces on $\mathcal{N}$ a classification in equivalence classes $[a, a' \in \mathcal{N}$ belong to the same class iff $a\rho a'$, and $a'\rho a]$. Now, ρ also induces a strict order on the set of classes. By construction, the Φ^{ρ} function is as follows: Given a problem (N, s), we order the agents in N into classes, according to by their types. Let us assume that we have $\#J$ different types and $k_1, ..., k_r$, $r \leq \#J$ different classes. Suppose that $k_1\rho k_2\rho ...\rho k_r$. Let $N_1 = \{a \in N : a \in k_1\}, ..., N_r = \{a \in N : a \in k_r\}$, and let $n_1 = \#N_1, ..., n_r = \#N_r$. We have that $n_1 + \cdots + n_r = n = \#N$. If $s < n_1$, then agents outside N_1 are definately not served. For agents in N_1, they all have the same chances to be served, namely, s/n_1. If an agent leaves, then there are no changes in the chances to be served of anybody else, whenever the leaving agent is outside N_1. Otherwise, agents in N_1 will increase their chances to $s/(n_1 - 1)$, while all others continue having no chances. If $n_1 \leq s < n_1 + n_2$, agents in N_1 are definately served for sure, agents in $N_3 \cup ... \cup N_r$ are definately not served, and agents in N_2 all have the same chances of being served, namely, $(s - n_1)/n_2$. Now, if an agent leaves, then there are no changes in the chances of being served for anybody else, whenever the leaving agent is in $N_3 \cup ... \cup N_r$. If it belongs to $N_1 \cup N_2$, agents outside N_2 do not change their right, while agents in N_2 will increase it to $(s - n_1)/(n_2 - 1)$. Similarly, we get pop-mon for all of the cases.

$\Longrightarrow$) Now let Φ be an anonymous and pop-monotonic allocation criterion. Define a binary relation on the set of potential agents, in the following way: For all $a, b \in \mathcal{N}$, $a\rho b \Leftrightarrow \Phi(\{a, b\}, 1) = \{a\}$ or $\Phi(\{a, b\}, 1) = \{\{a\}, \{b\}\}$. Let us see that ρ induces a priority relation on the set of types J.

First, the binary relation ρ on $\mathcal{N}$ is complete: for all $a, b \in \mathcal{N}$, either $a\rho b$, $b\rho a$ or both. Furthermore, it is transitive: If $a\rho b$, and $b\rho c$, then $a\rho c$. Suppose not. Then $\Phi(\{a,c\},1) = \{c\}$. But if this is the case, Φ is not an allocation criterion. Actually, there is no way of finding $\Phi(\{a,b,c\};1)$. It cannot be $\Phi(\{a,b,c\};1) = \{a\}$, since pop-mon would be violated: agent $a's$ right decreases when b leaves. Similarly, It cannot be $\Phi(\{a,b,c\};1) = \{b\}$, since agent $b's$ right decreases when c leaves. Also, it cannot be $\Phi(\{a,b,c\};1) = \{c\}$, since agent $c's$ right decreases when a leaves. It cannot be $\Phi(\{a,b,c\};1) = \{\{a\},\{b\}\}$, since agent $a's$ right decreases when b leaves, and it cannot be $\Phi(\{a,b,c\};1) = \{\{a\},\{c\}\}$, since agent $a's$ right decreases when b leaves. It cannot be $\Phi(\{a,b,c\};1) = \{\{b\},\{c\}\}$, since agent $b's$ right decreases when c leaves. Also, $\Phi(\{a,b,c\};1) \neq \{\{a\},\{b\},\{c\}\}$, since otherwise, agent $a's$ right decreases when b leaves.

Finally, note that ρ induces a priority relation on J, since Φ is anonymous. $\blacksquare$

Separability: Given $\{t,t'\} \in T$ and $T^c = T \backslash \{t,t'\}$, if we consider the types χ_1, χ_2, χ_3 and χ_4, such that,

$$(x_{1t}, x_{1t'}) = (x_{3t}, x_{3t'}) \neq (x_{2t}, x_{2t'}) = (x_{4it}, x_{4t'})$$

$$((x_{2t''})_{t''\in T^c}) = (x_{2t''})_{t''\in T^c} \neq (x_{3t''})_{t''\in T^c} = (x_{4t''})_{t''\in T^c},$$

it happens that $\chi_1\rho\chi_2 \Longleftrightarrow \chi_3\rho\chi_4$, for all $t, t' \in T$ and for all χ_1, χ_2, χ_3 and χ_4.

Theorem 2 (Gorman)[11]**:** *A weak order ρ on the set of types J can be represented by a point system iff ρ is separable in the characteristics.*

Appendix B

Visual incapacity levels	Activities the patient is unable to perform
Mild incapacity	Driving Reading small print (telephone directory, medicine bottles, etc)
Moderate incapacity	(The above) Doing fine handwork (sewing, putting in nails, etc) Reading a newspaper or book / filling out forms
Severe incapacity	(The above) Reading traffic, street or store signs Watching televison / seeing stairs Leisure activities such as bowling, gardening or window-shopping Playing cards, dominoes or bingo
Very severe incapacity	(The above) Cooking Reading large print in a book or newspaper Recognizing people at close range

Table1. Rank-ordered logit model —general

Variables	Levels	Parameters (t-ratio)	RI
Constant		-2.191 (-18.927)	
Visual incapacity	Mild	0.000	37.0%
	Moderate	0.718 (8.415)	
	Severe	1.540 (18.197)	
	Very severe	1.775 (20.922)	
Limitations on activities	Has some problems	0.000	16.1%
	Is unable	0.770 (13.001)	
Likelihood of improvement	50%	0.000	6.6%
	75%	0.220 (2.631)	
	99%	0.317 (4.397)	
Patient age (years)	50	1.376 (16.111)	28.7%
	65	0.877 (10.385)	
	75	0.543 (6.362)	
	85	0.000	
Waiting time (months)	3	0.000	11.6%
	6	0.324 (3.801)	
	12	0.512 (6.044)	
	18	0.557 (6.616)	
Log–likelihood function:-2644.895; CAIC:5360; n:100			

Table 2. Rank-ordered logit model —adjusted

Variables	Levels	Parameters (t–ratio)
Visual incapacity	Mild	0.000
	Moderate	0.700 (8.295)
	Severe	1.519 (18.127)
	Very severe	1.748 (20.804)
Limitations on activities	Has some problems	0.000
	Is unable	0.763 (12.925)
Ln (likelihood of improvement)		0.329 (8.256)
Ln (patient age (years))		−0.039 (−18.021)
Ln (waiting time (months))		0.297 (7.013)
Log–likelihood function:-2650.034; CAIC:5339; n:100		

References

[1] Ubel, P.A. Loewenstein, G. et al. Individual Utilities are inconsistent with rationing choices. *Medical Decision Making* 1996; 16: 108-116.

[2] Fricker J. BMA proposes strategy to reformulate waiting lists. *British Medical Journal* 1999; 318:78

[3] Elster, J. *Local justice* 1992. Sage Foundation cop. New York

[4] Pierce, G.A. et al. UNOS scientific registry: present and future. *Transplantation Proceedings* 1996; 28(4): 2355-2357.

[5] Leffell, M.S. et al. Impact of changes in the UNOS renal allocation system. *Transplantation Proceedings* 1997; 29(8):3435-3436.

[6] Sendom, M.E., J.K. French, D.J. Almos, K. Ramanathan et al.Waiting times and prioritisation for coronary artery bypass surgery in New Zealand. *Heart* 1999; 81:586-592.

[7] Jackson, N.W.,M.P. Doogue, J.M. Elliot. Priority points and cardiac events while waiting for coronary bypass surgery. *Heart* 1999; 81:367-373.

[8] Young, P. *Equity: in Theory and Practice* 1994. Princeton University Press 1994; 169-176.

[9] Debreu, G. Topological methods in cardinal utility theory. In: K. J. Arrow, S. Karlin and P. Suppes, eds. *Mathematical methods in the social sciences.* Stanford University Press, 1960; 16–26.

[10] Thomson, W. Population monotonic allocation rules. In: W.A. Barnett, H. Moulin, M. Salles and N. Schoffield, eds. *Social Choice, Welfare and Ethics.* Cambridge University Press, 1995; 79-124.

[11] Gorman, T. The Structure of utility functions. *Review of Economic Studies* 1968; 35:357-390.

[12] Green, P., V. Srinivasan. Conjoint analysis in consumer research: issues and outlook. *Journal of Consumer Research* 1978; 5:103-122.

[13] Ryan, M., J. Hughes. Using conjoint analysis to assess women's preferences for miscarriage management. *Health Economics* 1997; 6:261-273.

[14] Chakraborty, G., G. Gaeth, M. Cunningham. Understanding consumers preferences for dental service. *Journal of Health Care Marketing* 1993; 21:48-58.

[15] Mordue, A., D.W. Parkin, C. Baxter, M. Steward. Thresholds for treatment in cataract surgery. *Journal of Public Health Medicine* 1994; 16(4):393-398.

[16] Halliwell, T. How fair is cataract prioritisation? *New Zealand Medical Journal* 1998; 23 (1111:1076): 405-407.

[17] Hadorn, D.C., A.C. Holmes. The New Zealand priority criteria project. Part 1: overview. *British Medical Journal* 1997: 314:131.

[18] Steinberg E.P., Tielsch, J.M., Schein, O.D., et al. The VF-14. An idex of functional impairment in patients with cataract. *Arch Ophthalmol* 1994; 112:630-638.

[19] Alonso M., Andersen T.F., Cassard S.D., et al. International applicability of the VF-14. An index of visual function in patients with cataracts. *Ophthalmology* 1997; 104: 799-807.

[20] Schein, O.D., E.P. Steinberg, S.D. Cassard, et al. Predictors of outcome in patients who underwent cataract surgery. Ophthalmology 1995; 102: 817-823.

[21] Mangione, C.M., E.J. Orav, M.G. Lawrence, et al. Prediction of visual function after cataract surgery. *Arch Ophthalmol* 1995; 113:1305-1311.

[22] Hahn, G.J., S.S. Shapiro. A catalog and computer program for the design and analysis of orthogonal symmetric and asymmetric fractional factorial experiments. Report 66–C–165, 1966. Schenectady, N.Y.: General Electric Research and Development Center.

[23] Beggs, S., S. Cardell, J. Hausman. Assessing the potential demand for electric cards. *Journal of Econometrics* 1981; 16:1-19.

[24] Chapman, R., R. Staelin. Exploiting rank ordered choice set data within the stochastic utility model. *Journal of Marketing Research* 1982; XIX:288-301.

[25] McFadden, D. A conditional Logit model of qualitative choice behavior, in *Frontiers in Econometrics* 1974; Paul Zarembka, ed. New York: Academic Press, 105-42.

[26] Rosko, M, M. DeVita, W.F. McKenna, et al. Strategic marketing applications of conjoint analysis: an HMO perspective. *Journal of Health Care Marketing* 1985; 5(4): 27-38.