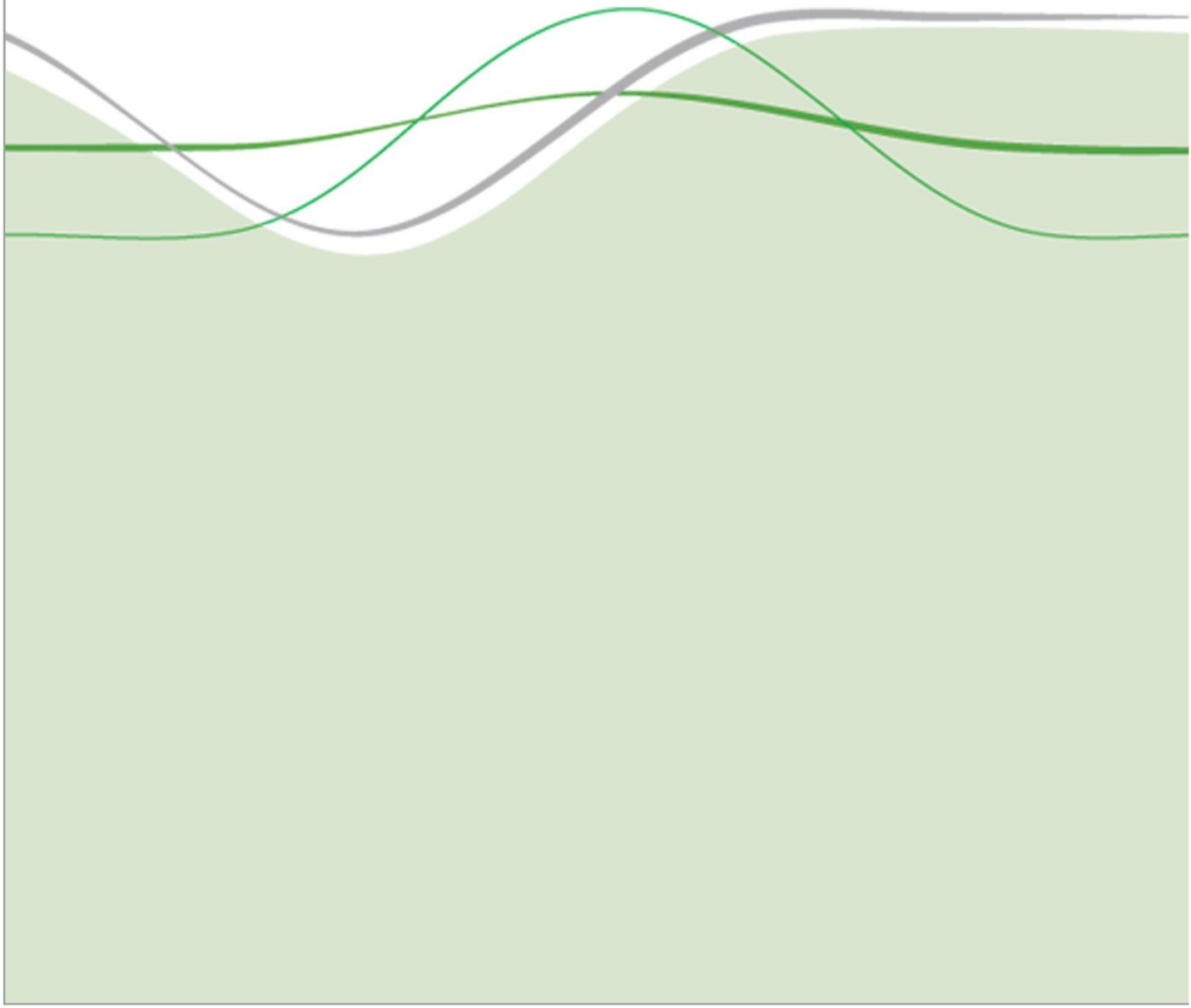
Office of Health Economics Research

Research Paper 18/03

New Methods for Analysing the Distribution of EQ-5D Observations

March 2018

Zamora B, Parkin D, Feng Y, Bateman A, Herdman M, Devlin N



New Methods for Analysing the Distribution of EQ-5D Observations

Zamora B[a], Parkin D[a,b], Feng Y[a], Bateman A[c], Herdman M[a], Devlin N[a]

[a]Office of Health Economics

[b]City, University of London

[c]Cambridgeshire Community Services NHS Trust and University of Cambridge

March 2018

For further information please contact:

Nancy Devlin

ndevlin@ohe.org

The Office of Health Economics (a company limited by guarantee of registered number 09848965)

Southside, 105 Victoria Street

London SW1E 6QT

United Kingdom

Tel: +44 207 747 8858

Table of Contents

ABSTRACT

Background: EQ-5D profile data are often under-analysed, but can yield important insights into levels of and changes in patient and population health. One characteristic is the extent to which they cluster together in a small number of profiles or are dispersed evenly over many profiles. This can have implications for interpreting statistical analysis of the corresponding EQ values (Index) data, and for clinical management of patients.

Aims: This paper aims to develop methods for describing observed distributions of EQ-5D profiles and to explore the properties of the new methods compared with existing ones (e.g. Shannon's Index). We investigate the methods using both real, from the English NHS, and simulated EQ-5D data, and show how they can be used to generate new insights into, for example, the differences between the three- and five-level versions of the EQ-5D with respect to how profile data are clustered.

Methods: We report three methods we have developed to characterise and summarise the distribution of health states in patient reported outcome (PRO) data within a sample or population of patients: the Health State Density Index (HSDI), Health State Density Curve (HSDC) and estimated Power Law functions (PLFs). We compare these and existing methods from information theory (e.g. Shannon's Index), in examining the distribution of EQ-5D health profiles in three data sets: across three groups of patients in Cambridgeshire Community Services NHS's electronic patient records for the EQ-5D-5L; the Health Survey for England 2014 for the EQ-5D-3L ; and the NHS PROMs programme for the EQ-5D-3L. The properties of the various methods are further examined using simulated data sets.

Results: Each method has different properties and will give different insights into patients' data. For example, the Shannon index (absolute and relative) is not sensitive to random variations but decreases slowly with "rare health states". The HSDI decreases slowly with random variations and is strongly affected by "rare" health states with large decreases towards zero (total inequality).

Conclusions: These methods can be used by researchers to better understand the characteristics of EQ-5D profile data. They can also be used by clinicians to understand the degree to which their patients' needs are homogeneous or characterised by distinct sub-groups, with implications for treatment planning. Finally, the methods can also be used as a way of comparing differences between instruments, such as the 3L and 5L, in measuring health.

Acknowledgements and disclaimers

This project was funded by a research grant from the EuroQol Research Foundation. All views expressed are those of the authors, and not necessarily those of the EuroQol Research Foundation.

The authors are grateful to Nils Gutacker, Thomas Kohlmann, and participants at the Health Economists' Study Group (HESG) Winter Meeting (Birmingham 2017) and the 34th EuroQol Plenary (Barcelona 2017), and to Nigel Rice, for helpful comments on earlier versions of this paper.

1. INTRODUCTION

EQ-5D data are collected for many purposes and used to inform many different types of decisions. Despite that, there is a tendency to focus analysis on EQ-5D values (also called the EQ Index[1]), essentially summarising the EQ-5D profile data collected from patients or populations using value sets as weights for levels within dimensions of the profile. There are several reasons why this may not be an adequate analytical approach.

First, the value sets are designed to meet the requirements of one particular application: the estimation of quality adjusted life years (QALYs) for cost effectiveness analysis. This is reflected both in the methods used to produce the value sets (stated preferences) and whose preferences are sought (by convention, the general public, rather than patients). Where the estimation of QALYs is not an aim of data collection, the rationale for using value sets to summarise profile data may be quite weak.

Secondly, value sets introduce an exogenous source of variance into the analysis of EQ-5D data. This can bias statistical inference (Parkin, Rice and Devlin, 2010). EQ values data may be used in assessing, for example, the effectiveness of a technology relative to its comparator in a clinical trial; whether there are statistically significant differences between regions in population health; or which hospitals provide better quality of care. In each case, the conclusions from statistical analysis may depend on the properties of the value set used. Different value sets, which place a greater or lesser weight on dimensions and levels and interactions between them, could yield different statistical results. More generally, there is no neutral way of summarising EQ-5D profile data. This is not a criticism of the EQ-5D or EQ-5D values; the same can be said of all weighted index approaches to generating summary scores of PRO instruments.

Thirdly, focusing on EQ-5D values means that important insights from patient data may be missed. EQ-5D profile data tell us in which dimensions problems are experienced, and with what magnitude – and for changes in health, which dimensions change and by how much. Focussing analysis on what is happening within each dimension provides information that is obscured when they are aggregated to EQ-5D values (Gutacker et al., 2013). The EQ-VAS tells us something additional – the patients' overall assessment of their own health, relating both to factors captured in the EQ-5D profile and other factors related to a broader underlying construct of health, including aspects not related to a patient's condition (Feng, Parkin and Devlin, 2014).

In previous work, we sought to understand the nature of distributions of EQ-5D values data, and noted that these are driven both by the properties of the value sets *and* distinctive distributions of EQ-5D profile data. For example, we noted that typical distribution of EQ-5D-3L values data arising from use of the value set published by the Measuring and Valuing Health (MVH) project at the University of York (Dolan, 1997) – characterised by two groups – arises from both the peculiarities of this value set and the fact that only certain profiles appear in these groups (Parkin, Devlin and Feng, 2016).

[1] We use the term 'EQ-5D values' rather than 'EQ Index' to refer to the summary measure of a profile created using a value set, to avoid confusion with the term index used elsewhere in this paper to refer to measures od distributional properties.

We also noted that, when examining EQ-5D-3L data across multiple, large data sets, a surprisingly small number of profiles (<12) can account for the vast majority (80%) of the data (Parkin et al., 2016). We also observed that often within particular conditions only some levels within each dimension are observed – which affects assessment of health improvements by giving very specific starting points. We have observed similar phenomena in analysis of EQ-5D-5L data – although more profiles are observed, and the data are less clustered (Feng et al., 2016). With respect to both the 3L and the 5L, some theoretically possible profiles, such as many problems on one dimension and no problems on all others, are very rarely observed.

This has led us to focus more on the nature of EQ-5D profile data. Examining cumulative frequencies is a simple but effective way of capturing important information about the distribution of these data, but does not provide a means of summarising and reporting the concentration of profiles in data sets in a manner that would allow us to compare and contrast this characteristic between different data sets.

This led us to think about ways of summarising the extent to which observations cluster on a small number of profiles or are a wide range. We recognised analogies to measures of the distribution of income, such as the Lorenz curve and Gini coefficient (Gini, 1921; Lorenz, 1905); with indices used to measure information complexity and species diversity, such as Shannon's index; and with power law analyses that relate the number of profiles to the frequency with which they are observed.

Such measures facilitate better understanding of the profile data underpinning EQ-5D values data. They enable comparisons between different PRO instruments, including different versions of the EQ-5D. They may also be relevant to clinical management. For example, homogenous patient groups can be managed by standardised treatment protocols, whereas dispersed EQ-5D profiles suggest a need for more complex, individualised patient management.

The aim of this study is to develop methods for describing observed distributions of EQ-5D profiles and to explore the properties of the new methods compared with existing measures. We demonstrate the use of the methods using both real EQ-5D data from the English NHS, and simulated EQ-5D data, and show how the methods can be used to generate new insights into, for example, the differences between the three- and five-level versions of the EQ-5D in terms of how profile data are clustered.

2. DATA

Cambridgeshire Community Services musculoskeletal therapy, specialist nursing, and community rehabilitation patients completed the EQ-5D-5L before treatment, with resulting data stored in NHS's electronic patient records data warehouse. There are 30,284 observations.

The 2014 Health Survey for England (HSE) comprised a multi-stage stratified sample representing the population in England living in private households. A complete EQ-5D-3L profile was available for 7,085 people aged 16 years and over.

The English NHS Patient Reported Outcome Measures programme (PROMs) collects EQ-5D-3L data before and after surgery from patients undergoing hip and knee replacements and varicose vein and groin hernia repairs in all providers of NHS-funded care in England. We used data from three years: 2009-2010, 2010-2011, and 2011-2012.

Table 1 presents the sample sizes for the data sets and statistics on age and gender. All data sets only include patients with complete EQ-5D profile data.

Table 1. Sample characteristics for the data sets

	N	% female	Mean age (years)	Age range
Cambridgeshire NHS: All Patients	30,284	59.6	59.1	13-104
Cambridgeshire NHS: MSK	19,999	59.8	52.5	13-96
Cambridgeshire NHS: Specialist Nursing	3,366	52.1	67.0	17-104
Cambridgeshire NHS: Rehabilitation	6,919	62.5	74.4	18-103
Health Survey for England 2014	7,085	56.1	49.7	16 and over
NHS PROMs Hernia 2009-10	19,416	1.3	62.8	20 and over
NHS PROMs Hernia 2010-11	21,265	1.2	62.7	20 and over
NHS PROMs Hernia 2011-12	7,927	0.1	64.8	20 and over
NHS PROMs Hips 2009-10	29,506	59.9	68.8	20 and over
NHS PROMs Hips 2010-11	37,923	60.6	68.6	20 and over
NHS PROMs Hips 2011-12	5,194	67.4	69.8	20 and over
NHS PROMs Knee 2009-10	32,078	56.7	69.7	20 and over
NHS PROMs Knee 2010-11	39,098	57.5	69.7	20 and over
NHS PROMs Knee 2011-12	4,913	61.3	70.1	20 and over
NHS PROMs Varicose Veins 2009-10	7,923	72.0	52.6	20 and over
NHS PROMs Varicose Veins 2010-11	8,141	71.4	53.0	20 and over
NHS PROMs Varicose Veins 2011-12	2,258	76.5	54.4	20 and over

Note: Mean age in the NHS PROMs data was estimated as the mean of the 10-years intervals reported in the data, where the mean was imputed from the corresponding intervals of age in HSE data.

3. METHODS

Building on our earlier research on EQ-5D distributions, we have developed new ways of describing the distributions of health states, based on existing indexes such as the Gini coefficient and the Shannon index. We use them to describe and summarise the degree of differentiation of people across EQ-5D health states, but they could just as readily be applied to other patient reported outcome (PRO) instruments.

3.1. Health State Density Curve and Health State Density Index

The evenness of the distribution of categorical health profiles data can be described using a *Health State Density Curve* (HSDC) and a *Health State Density Index* (HSDI). These are analogous to the Lorenz curve, a cumulative frequency curve that compares a real distribution of a variable with a uniform distribution representing equality, and its associated Gini coefficient.

The advantage of these measures is that they can be applied to the distribution of any variable, categorical or continuous, whilst other inequality indexes, such as Theil's entropy measures, are specifically designed for continuous variables such as income. However, the categorical variable must be ordered, and health profiles typically are not fully-ordered, as some are not comparable using prior logic (Devlin, Parkin and Browne, 2010). The distribution of health profiles in the HSDC are therefore ranked from the most to the least frequent profile, where the ranking of equally frequent profiles is arbitrary. The vector of cumulative distribution of the total observed distinct *S* health profiles y= $\left(\frac{1}{S}, \frac{2}{S}, .., \frac{S-1}{S}, 1\right)$ is represented in the Y-axis, and the cumulative distribution of patients $x_i = \sum_{j=1}^{i} \frac{N_j}{N}$ associated with each health profile j, from i=1 to i=S, is represented in the X-axis, where N_j is the number of observations in profile j and N is the sample size. The resulting graphical function is the HSDC.

The HSDI is based on the area between the diagonal of total equality and the HSDC. The HSDI has a range of [0,1]. Total equality in the distribution of profile data (that is, the case where all profiles account for an equal share of observations) would be shown as a 45°diagonal line on the HSDC and a HSDI of 1. Total inequality would lead to an HSDC with a ⌟ shape, and a HSDI of 0. . This is analogous to the definition of the Gini coefficient, and can be calculated as:

(1) $$HSDI = \sum_{i=1}^{S} (x_i - x_{i-1})(y_i + y_{i-1})$$

where (x_i, y_i) are respectively the points of the X- and Y-axis, representing the cumulative distribution of patients and the cumulative distribution of profiles, as calculated at each distinct profile i=1,2,..,S, ranked from i=1 being the most frequent profile to i=S being the least frequent profile.[2]

[2] The HSDI for a unique profile is: $HSDI = x_1 y_1 = 1$, since both cumulative proportions of patients and profiles are 1 in a unique profile. The product $x_1 y_1$ is also the first summand in formula (1) for the first (most frequent) profile in distributions with if S>1.

The HSDI has several properties as an inequality or evenness index:

(a) Independence of number of observed health profiles (S). Following the method proposed in Smith and Wilson (1996), we tested the values of HSDI for 1,2,3,4,5,10,20, and 40 repetitions of a distribution of observations of 4 health profiles with (1479, 1, 1, 1) observations each, so that S varies between 4, for one repetition, and 160, for 40 repetitions. The HSDI has the same value of 0.25 for the 8 replications, so it is independent of *S*.

(b) The range of HSDI is the unit interval. HSDI reaches one for total equality, which is the case where there are the same number of patients in all health profiles, and HSDI is zero for total inequality. The case of a unique health profile results in total equality since empty profiles do modify the value of the HSDI index. We illustrate below the case of total inequality as depending on the number of rare profiles (profiles with very low number of observations relative to the most frequent profiles).

(c) The HSDI changes linearly with linear changes in the distribution of patients. We calculated the HSDI for the Molinari sequence (Molinari, 1989) which distributes 1000 observations between two profiles from total inequality (999,1) to total equality (500,500) by switching 99/100 observations in five steps. The results of the HSDI in this sequence go from a minimum of 0.5 (for the most unequal distribution 999-1) to HSDI=1 (for the total equality pair 500-500), increasing by 0.1 in five equal steps.

3.2. Shannon Index and Shannon Evenness

The Shannon Index (H') was derived as a measure of information content (Shannon, 1948) but many of its properties and extensions have been developed to measure biodiversity. It is calculated as:

(2) $$H' = -\sum_{i=1}^{S} p_i \log(p_i)$$

where p_i denotes the proportion of patients in health profile i, and *S* is the number of observed distinct profiles. In most cases, especially in information theory, the Shannon entropy index is calculated using the natural logarithm, which is what we use. However, some applications of the Shannon Index in health-related quality of life (HRQoL) research (Agborsangaya et al., 2014; Janssen et al., 2013; Von Steinbuechel et al., 2016) defined the index using logarithm base 2. In these studies H' has been interpreted as capturing the discrimination properties of HRQoL instruments, with larger values of H' indicating more absolute discriminatory power of the instrument.

In general, information or diversity contained in H' measured from samples is lower than the true maximum population or community diversity H, which generates a bias in the Shannon index and an unwanted dependency on the number of different variants in a sample, S. As Pielou (1966) remarks, H' cannot be regarded as exactly equal to the true diversity H even when the population sizes of each profile, N_i, are exactly known, and then $p_i = \frac{N_i}{N}$, unless the observed population is a sample from a conceptually infinite "super-population" of which it is exactly representative.

The expected value of (2) when $p_i = \frac{N_i}{N}$ is calculated using natural logarithms and neglecting terms of order N^{-2} (Lande, 1996; Pielou, 1966) is:

(3) $$E(H') = H - \frac{\bar{S}-1}{2N}$$

where $\bar{S}$ is the upper limit of number of health profiles as if obtained from an infinite population. This is generally unknown in ecological samples, but the EQ-5D instruments define the maximum number of health profiles: $\bar{S} = 3^5 = 243$ for the 3L, and $\bar{S} = 5^5 = 3{,}125$ in the 5L. Therefore, for any sample of size N, the bias of the Shannon index (Bias H') can be measured as:

(4) $$Bias\ H' = \frac{\bar{S}-1}{2N}$$

Therefore, the Shannon Index is underestimated in a sample of size N, and the bias is larger for the 5L than for the 3L. This bias ensures that the 5L is larger or has more absolute discriminatory power than the 3L, even if they have the same H'; for example, if the number of observed health profiles is the same in both instruments (S) and patients are uniformly distributed across profiles, H'=log(S). For a given sample size, we can measure this theoretical difference in absolute discriminatory power as (3,125-243)/2N, which is the difference in the bias between the instruments. Some existing comparisons of the 3L and 5L do not measure or acknowledge the bias when ascertaining a larger discriminatory power of the 5L than of the 3L (e.g. Janssen *et al*, 2012). For samples of more than 1000 patients, the bias can be negligible for the 3L but it should not be ignored when interpreting 5L results. Notice that this bias only depends on the theoretical maximum number of profiles (e.g. $\bar{S} = 243$, for the 3L, and $\bar{S} = 3{,}125$ for the 5L) the sample size but not on the number of observed distinct profiles.

Pielou (1975) defined an evenness index based on the Shannon Index as:

(5) $$J' = \frac{H'}{\ln(S)}$$

The J' index is interpreted as relative discriminatory power since it is defined as the Shannon Index H' (absolute discriminatory power) divided by its maximum, log(S), which is the maximum absolute discriminatory power as obtained from the uniform distribution, that is, with the same number of persons in each health profile. The J' index is intended to compensate for the dependency of H' on S so that an increase in information content of the index is the result of observing relatively more patients in each profile (increase in p_i), and not just a result of increasing the number, S, of populated health profiles. In summary, the properties of the Shannon indexes are:

(a) Independence of number of observed distinct health profiles (S). A property of information content is that a system contains more information the more possible states it contains. The Shannon index measures this, as it is an increasing monotonic function of S. A second property is that there is more information when the probability of encountering each state is high, e.g. all states are equally abundant. The J' index (relative discriminatory power) captures this, which requires J' to be independent of S. As Smith and Wilson (1996) show, the evenness index J' is not independent of S

for low values of S if measured from an uneven distribution, but it becomes asymptotically stable after S reaches about 25 if this distribution is replicated.

(b) The range of H' is [0, log(S)] and the range of J' is [0,1]. H' and J' reach maximum value, log(S) and one, respectively, for total equality or maximum discriminatory power, which is the case where there are the same number of patients in all health profiles.

(c) The H' and J' indexes have the Molinari (parabolic) shape. We have calculated the H' and J' indexes for the Molinari sequence, including 11 communities distributing 1000 patients between two profiles. Both indexes increase from zero at the pair of profiles 999-1 to their maximum for the equal distribution 500-500. The indexes grow following a concave shape such that the rate of growth decreases when approaching total equality.

3.3. Comparing the HSDI and Shannon Evenness J'

We first compare the indexes in terms of information theory or discriminatory power, where larger values of the index imply a more even distribution of patients across profiles, associated with greater differentiation between groups of patients. For example, an index close to zero means that a large proportion of patients are described by just one profile. In this case, we can accurately predict the health profile of a patient without needing any patient information. We would like also to know the sensitivity of the indexes to random changes in the number of patients which do not reflect changes in the underlying evenness of the population.

Secondly, we compare how sensitive the indexes are to rare versus abundant profiles. The ecological literature describes how the Shannon index is "strongly affected by importance of species in the middle of the sequence. For large samples, H' is consequently somewhat damped against effects of differences in quantitative proportions of the first few species. Effect of the rarer species are also damped" (Whittaker, 1972). These properties mean that H' and J' indexes do not capture properly the EQ-5D typical abundance of rare profiles. This will be shown when comparing J' with HSDI. Firstly, we present both indexes calculated from the Molinari sequence in Figure 1. The HSDI changes linearly and it does not achieve the minimum of total inequality at zero. This is a consequence of having only two profiles (S=2); adding rare profiles of 1 patient to the sequence with 999 patients in the most abundant profile would allow to achieve HSDI=0.06 at S=30.

In addition of the simulations with the Molinari sequence, we present simulations of the EQ-5D-3L and EQ-5D-5L profiles from the multinomial distribution, starting from large random samples (N=25,000 for the 5L and N=10,000 for the 3L) of the Dirichlet distribution reflecting total evenness, that is, identical prior probabilities of each level for each of the five dimensions. This uniform distribution of health profiles is rarely observed for the EQ-5D in patient populations. We also construct uneven distributions from the Dirichlet distribution with different number of observed profiles. Our approach to study the index properties from these simulated samples is by computing

the indexes for each new health profile within each sample instead of by drawing new random samples. This allows to test the evolution of the indexes with different values of profile concentration p_i. Therefore, for each simulated sample, we compute the changes in the indexes for each additional profile where profiles are ranked from 1 to S, from the most to the least frequently observed.

Then, for each simulated distribution we obtain a sequence of sequence of indexes from profile 1 to S, HSDIs and J's. We have observed a decreasing linear relationship between each of these sequences, HSDIs and J's, and the proportion of rare profiles. We present an approximation to the proportion of rare profiles which is specified below and only depends on p_i, for i=1,..,S. A large proportion of rare profiles captures large differences across p_i, since rare profiles are those with very low values for p_i placed at the latest positions in the sequence.

The estimation of the proportion of rare profiles we have considered comes from the fact that the number of rare profiles (those with very few patients) can be (upwards) approximated as the difference between the number of observed profiles, S, and the numbers equivalent which is a concept used in ecological diversity measured by the exponential of the Shannon index H'. The numbers equivalent of a diversity index is the number of equally likely elements needed to produce the given value of the diversity index (Jost, 2007).

By calculating the series of $HSDI_s$ and J'_s indexes for each sample, setting the sequence of the subscript s from s=1 to s=S, with HSDI=J'=1 when s=1, we estimate the following:

(6) $$HSDI_s = \alpha + \beta(proportion\ rare\ profiles)_s + \varepsilon_s$$

(7) $$J'_s = \gamma + \delta(proportion\ rare\ profiles)_s + \varepsilon_s$$

Where the explanatory variable proportion of rare profiles is estimated as:

(8) $$(proportion\ rare\ profiles)_s = \frac{s - exp(H'_s)}{s},\ with\ s = 1, 2, .., S$$

The proportion of rare profiles is zero for the first profile and for the uniform distribution since the Shannon index takes the values $H'_1 = 0, H'_s = \log(s)\ if\ p_i = \bar{p}\ for\ all\ i = 1, .., s.$

Table 2 presents results from the linear regressions of both relative indexes, HSDI and J', on the proportion of rare profiles obtained from different simulated samples of the Dirichlet distribution. Table 2 also includes the maximum number of observed profiles S, the maximum percentage of rare profiles at s=S, and the values of the indexes for the first two profiles s=2, and the complete sample at s=S.

Table 2. Linear regression of HSDI and J' on proportion of rare profiles

	Simulations EQ-5D-5L profiles							
	Total evenness		Uneven 230 profiles		Uneven 40 profiles		Uneven 40 profiles	
	HSDI	J	HSDI	J	HSDI	J	HSDI	J
Slope	-2.9182	-0.1237	-0.9153	-0.3035	-2.1743	-0.2691	-1.3839	-0.3101
Intercept	0.9542	0.9998	0.8418	1.0030	0.9403	0.9976	0.9149	0.9984
R2	0.9586	0.9999	0.9766	0.9946	0.9520	0.9971	0.9214	0.9989
S	3124	3124	230	230	40	40	40	40
Index (at s=2, s=S)	(1,0.80)	(1,0.99)	(0.97,0.23)	(0.99,0.78)	(0.97,0.76)	(0.99,0.97)	(0.97,0.56)	(0.99,0.90)
proportion rare (at s=S)	0.062		0.696		0.105		0.297	
(min.- max) obs per profile	(1-17)		(1-1329)		(286,1461)		(1-1766)	
N (total obs.)	25000		25000		25000		25000	
	Simulations EQ-5D-3L profiles							
	Total evenness		Uneven 34 profiles		Uneven 22 profiles		Uneven 12 profiles	
	HSDI	J	HSDI	J	HSDI	J	HSDI	J
Slope	-6.0559	-0.1737	-0.8907	-0.5915	-1.0337	-0.5492	-1.0253	-0.5881
intercept	0.9778	0.9999	0.8721	0.9078	0.9238	0.9963	0.8960	0.9894
R2	0.9557	0.9995	0.9796	0.9706	0.9786	0.9992	0.9662	0.9987
S	243	243	34	34	22	22	12	12
Index (at s=2, s=S)	(0.99,0.90)	(0.99,0.99)	(0.71,0.09)	(0.74,0.37)	(0.99,0.23)	(0.99,0.60)	(0.90,0.32)	(0.97,0.64)
proportion rare (at s=S)	0.014		0.891		0.707		0.592	
(min.- max) obs per profile	(21-59)		(1-6322)		(1-2481)		(1-3609)	
N (total obs.)	10000		10000		10000		10000	

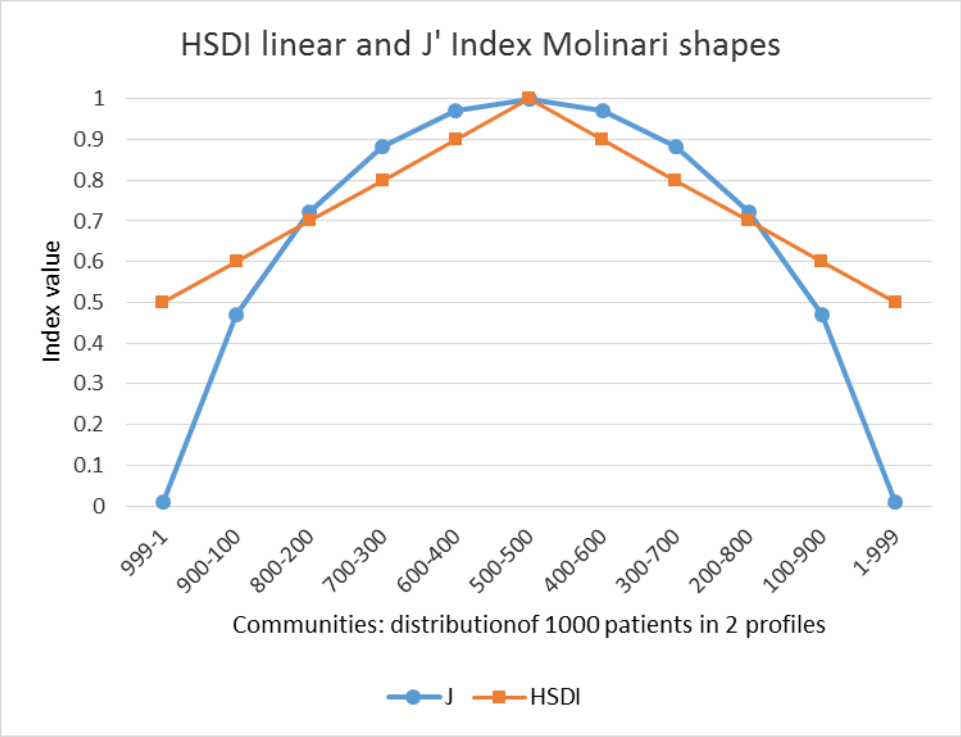
Notes:
All regression coefficients are statistically significant with p-value<0.01
In the simulations of total evenness, the differences in numbers of observations per profile are random.

The first differential characteristic observed refers to the sensitivity to random variations with no rare profiles, as resulting from the "total evenness" simulations. In this case, the decrease in J' cannot be perceived given the low proportion of rare observations, but HSDI is very sensitive to these random variations as if they were rare observations; HSDI decreases from 1 to 0.8. This behaviour of HSDI resulting from random variations cannot be distinguished with that from an uneven sample of 40 profiles with no rare profiles (proportion of rare profiles around 0.1) whilst J' is slightly more sensitive to capture this underlying heterogeneity than random variations and decreases from 0.99 to 0.97.

The second differential characteristic refers to the sensitivity to rare profiles, with the slope of the HSDI approaching -1 which would render a zero value for the HSDI for a proportion of rare observations equal to one. In contrast, the estimated slopes for J' in the simulations with rare observations point out a slower decrease of J' with the addition of rare profiles. Comparing the slope from the regressions from J', a robust finding is that J's is more sensitive to rare profiles for the 3L than for the 5L.

Therefore, J' indicates larger relative discriminatory power than HSDI, and this is not affected by random variations in the importance of profiles. HSDI informs better than J' on the effect and importance of rare profiles.

Figure 1. HSDI and J in Molinari sequence



3.4. Power Law Analysis

Another way to express the relationship between profiles and observations in a data set is to estimate its functional form as a power law relationship. This is described by Ball (2005) as: "if the value of some quantity y depends on the value of another quantity x according to a power law relationship, this means that each time x is doubled, y increases by some constant factor. The exponent of the power law is a number that tells us how big this factor is." A power law is a special kind of probability distribution which has been associated to inequality, for example power law behaviour with a Pareto distribution is observed in size distributions of economic interest such as cities, firms, income, and wealth.

Specifically, we can link the HSDC with a Pareto distribution or power law for which the Lorenz curve takes the form:

(9) $$y_i = 1 - (1 - x_i)^{\frac{(\alpha-1)}{\alpha}}, with\ \alpha > 1$$

The parameter α increases when approaching total evenness, or the diagonal of the HSDC. According to the expression of the Gini coefficient for this distribution, the value of the HSDI can be estimated as:

(10) $$HSDI = 1 - \frac{1}{2\alpha - 1}$$

Therefore, it is possible to approximate the cumulative distribution of EQ-5D profiles by a Pareto distribution of parameter α according to the value of the HSDI.

If we approximate the HSDC in (9) in semi-logarithmic form so that we predict the cumulative proportion of observations from the cumulative proportion of profiles, we can estimate the following regression:

(11) $$x_i = a + b\log(y_i) + \varepsilon$$

where y_i is the cumulative number of profiles, x_i is the cumulative number of observations up to the i[th] profile and ε is random error. A potential use of this is to predict the number of profiles in a sample of a given size. The intercept corresponds to the point $y_S = 1$, so $a = 1$.

We have estimated the semi-logarithmic function in equation (11) and the power law for our simulated samples of the EQ-5D profiles and present the results in Table 3 which can be interpreted in relation to those of the HSDI and J indexes presented in Table 2.

Table 3. Estimations of the power law from simulated samples

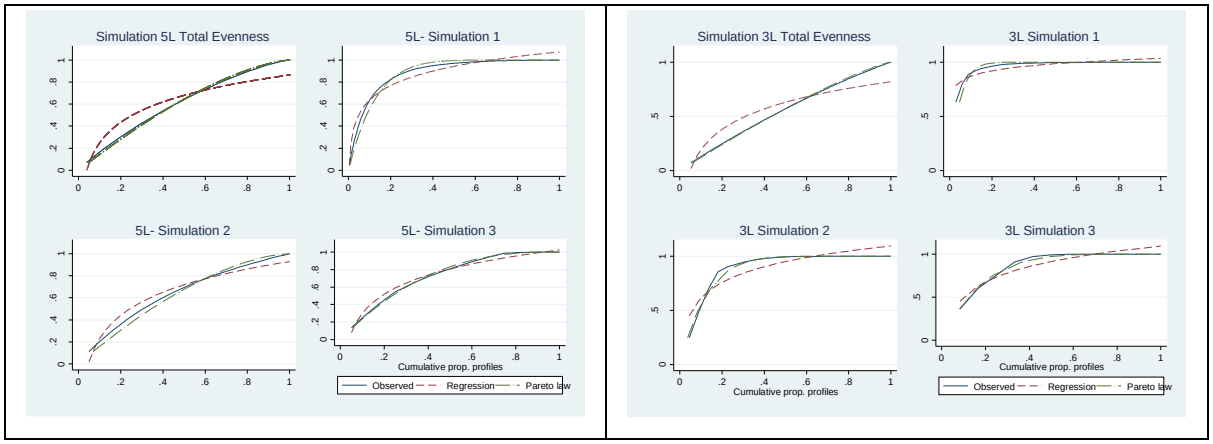
	Simulations of EQ-5D-5L profiles				**Simulations of EQ-5D-3L profiles**			
	Total evenness	Simula-tion 1	Simula-tion 2	Simula-tion 3	Total evenness	Simula-tion 1	Simula-tion 2	Simula-tion 3
log(yt)	0.2656	0.1900	0.3037	0.3143	0.2735	0.0716	0.2088	0.2570
intercept	0.8648	1.0743	0.9285	1.0249	0.8194	1.0353	1.0946	1.0926
R^2	0.8367	0.9312	0.9432	0.9679	0.8233	0.7294	0.8052	0.8848
S	3124	230	40	40	243	34	22	12
proportion rare (at s=S)	0.062	0.696	0.105	0.297	0.014	0.891	0.707	0.592
(min.- max) obs per profile	(1-17)	(1-1329)	(286-1461)	(1-1766)	(21-59)	(1-6322)	(1-2481)	(1-3609)
N (total obs.)	25000	25000	25000	25000	10000	10000	10000	10000
HSDI	0.8	0.23	0.76	0.56	0.9	0.09	0.23	0.32
Pareto law α	3	1.15	2.58	1.64	5.5	1.05	1.15	1.24

The goodness of fit of the power law, captured by R-squared, is better for samples with proportion of rare profiles in the middle ranges than for uniform samples or with a very large proportion of rare profiles. Moreover, for uniform samples the intercept is around 0.8 meaning an underestimation of the cumulative percentage of observations when considering all profiles. The slope of the power law (the exponent) is smaller for samples with large proportions of rare profiles, meaning a prediction of a larger cumulative proportion of observations for any given proportion of profiles than in other samples. This is consistent with lower values of the HSDI and J' indexes, and with the associated lower value of the parameter $\hat{\alpha}$ of the Pareto power law as derived from (10) which is presented at

the bottom row in Table 3. Remember that the power law parameter α increases when approaching total evenness, or the diagonal of the HSDC.

Figure 2 represents the power law comparing the observed relationship with the predicted values from regression (11) and with the Pareto law distribution of parameter $\widehat{\alpha}$. The distribution of simulated samples fits better the Pareto law than the semi-logarithmic regression (11), but both functions approach with the typical concentrated distribution of EQ-5D data.

Figure 2: Power Law distributions



4. RESULTS

4.1. Distributional characteristics of EQ-5D profiles

Tables 4, 5, and 6 summarise key characteristics of the EQ-5D data from the different data sets. The characteristics of the PROMs data are almost constant over the three years, so this summary is for a pooled data set.

Table 4. Distribution of EQ-5D dimensions and levels

	Mobility	Self-care	Usual Activities	Pain/ Discomfort	Anxiety/ Depression
Cambridgeshire NHS All Patients					
No problems	31.25	54.54	17.57	10.86	52.73
Slight problems	25.54	23.86	28.84	26.94	25.70
Moderate problems	27.48	15.29	30.63	39.35	15.54
Severe problems	13.42	4.30	13.29	18.94	4.20
Extreme problems	2.31	2.00	9.67	3.90	1.83
Cambridgeshire NHS MSK					
No problems	37.07	61.69	16.05	3.63	56.16
Slight problems	26.96	23.00	33.32	26.74	24.40
Moderate problems	24.44	12.23	32.81	43.70	14.01
Severe problems	10.83	2.64	12.96	21.70	3.89
Extreme problems	0.71	0.45	4.87	4.24	1.54
Cambridgeshire NHS Specialist Nursing					
No problems	33.10	57.40	35.68	37.40	46.79
Slight problems	19.88	17.53	22.01	28.55	28.64
Moderate problems	24.81	14.32	22.58	22.67	17.80
Severe problems	17.80	6.24	12.03	9.27	4.63
Extreme problems	4.43	4.52	7.69	2.11	2.14
Cambridgeshire NHS Rehabilitation					
No problems	13.53	32.49	13.17	18.88	45.70
Slight problems	24.19	29.46	19.21	26.74	28.02
Moderate problems	37.56	24.61	28.26	34.92	18.88
Severe problems	18.79	8.17	14.87	15.67	4.89
Extreme problems	5.93	5.28	24.50	3.80	2.51
Health Survey for England 2014					
No problems	82.67	94.52	84.18	67.20	80.76
Moderate problems	17.15	5.17	14.42	28.98	17.06
Extreme problems	0.18	0.31	1.40	3.82	2.17
NHS PROMs Hernia-Before					
No problems	79.28	96.40	71.41	32.23	84.25
Moderate problems	20.65	3.44	26.57	63.75	14.80
Extreme problems	0.07	0.15	2.02	4.02	0.96
NHS PROMs Hips-Before					
No problems	6.40	44.84	6.16	0.94	57.52
Moderate problems	93.17	53.96	74.20	57.40	37.58
Extreme problems	0.43	1.20	19.63	41.66	4.90
NHS PROMs Knee-Before					
No problems	6.25	68.47	8.86	0.99	62.29
Moderate problems	93.46	30.78	77.41	59.50	33.90
Extreme problems	0.29	0.75	13.73	39.51	3.81
NHS PROMs Varicose Vein-Before					
No problems	77.39	96.74	75.93	26.88	78.43
Moderate problems	22.50	3.12	22.91	67.10	19.65
Extreme problems	0.11	0.14	1.16	6.03	1.93
NHS PROMs Hernia-After					
No problems	82.80	95.07	78.83	66.47	87.44
Moderate problems	17.13	4.73	19.74	31.74	11.71
Extreme problems	0.07	0.21	1.43	1.79	0.85

Table 5 (cont.). Distribution of EQ-5D dimensions and levels

	Mobility	Self-care	Usual Activities	Pain/ Discomfort	Anxiety/ Depression
NHS PROMs Hips-After					
No problems	54.75	76.33	50.60	51.79	80.36
Moderate problems	45.14	23.00	44.96	44.17	17.67
Extreme problems	0.11	0.66	4.44	4.04	1.97
NHS PROMs Knee-After					
No problems	46.48	78.25	41.73	32.09	75.97
Moderate problems	53.38	21.16	52.93	61.50	21.52
Extreme problems	0.13	0.59	5.35	6.41	2.51
NHS PROMs Varicose Vein-After					
No problems	82.91	96.20	83.25	61.15	83.47
Moderate problems	17.05	3.70	15.72	35.80	14.90
Extreme problems	0.05	0.10	1.03	3.05	1.63

The HSE covers the general population, and is on average younger than our patient population, and might be expected to be the healthiest. However, Table 4 suggests that Varicose Veins and Groin Hernia patients have even better health states after surgery than the general population. Conversely, the health profiles of the three oldest groups, Rehabilitation and Hip and Knee patients, have many more problems.

Table 6. 10 most frequent EQ-5D-5L profiles

All Patients		**MSK Patients**		**Specialist Nursing**		**Rehabilitation**	
Profile	Cumul. %	Profile	Cumul. %	Profile	Cumul. %	Profile	Cumul. %
11121	0.0397	11121	0.0495	11111	0.1319	11111	0.0186
11111	0.0717	11221	0.0948	11121	0.1714	21221	0.0304
11221	0.1036	21221	0.1282	11112	0.2056	11121	0.0416
21221	0.1297	21231	0.1602	11122	0.2219	32331	0.0519
21231	0.1527	11231	0.1869	21121	0.2368	33331	0.0616
11231	0.1712	11131	0.2091	11113	0.2499	32332	0.0711
31331	0.1884	31331	0.2312	21221	0.2614	22221	0.0795
11131	0.2048	11111	0.2510	21111	0.2727	21111	0.0879
11331	0.2163	11331	0.2677	11211	0.2834	33332	0.0954
21331	0.2277	21331	0.2837	21211	0.2938	21121	0.1028

There is less concentration in Rehabilitation patients than the other two groups, with 10% compared with almost 30% in the 10 most frequent profiles.

Table 7: 10 most frequent EQ-5D-3L profiles

HSE		NHS PROMs							
		Groin Hernia-Before		Groin Hernia-After		Hip Replacement-Before		Hip Replacement-After	
Profile	Cumul. %	Profile	Cumul. %	Profile	Cumul. %	Profile	Cumul. %	Profile	Cumul. %
11111	0.5743	11121	0.3075	11111	0.5688	21221	0.1664	11111	0.3556
11121	0.6829	11111	0.5786	11121	0.6978	22221	0.2887	21221	0.4552
11112	0.7485	11221	0.6711	11221	0.7415	22232	0.3757	22221	0.5260
11122	0.7812	21221	0.7374	21221	0.7839	22222	0.4551	11121	0.5823
21221	0.8135	11122	0.7815	11112	0.8136	22231	0.5310	22222	0.6361
21121	0.8438	21121	0.8174	21222	0.8316	21222	0.5938	11211	0.6721
21222	0.8611	21222	0.8414	21111	0.8493	22332	0.6538	21222	0.7076
11221	0.8738	11112	0.8629	11122	0.8664	21231	0.7057	11221	0.7428
22221	0.8840	11222	0.8827	21121	0.8811	22331	0.7396	21211	0.7762
22222	0.8933	22221	0.8929	21211	0.8956	21232	0.7705	21121	0.8008

Table 6 (cont.): 10 most frequent EQ-5D-3L profiles

NHS PROMs							
Knee Replace. -Before		Knee Replacement-After		Varicose Vein-Before		Varicose Vein-After	
Profile	Cumul. %	Profile	Cumul. %	Profile	Cumul. %	Profile	Cumul. %
21221	0.2524	11111	0.2246	11121	0.3538	11111	0.5426
21231	0.3543	21221	0.3789	11111	0.5771	11121	0.7030
21222	0.4439	11121	0.4835	11122	0.6464	21221	0.7435
22232	0.5052	22221	0.5483	21221	0.7042	11112	0.7829
22221	0.5637	21222	0.6090	11221	0.7558	11122	0.8192
21232	0.6207	11221	0.6679	21121	0.7992	21222	0.8453
22231	0.6684	22222	0.7267	21222	0.8314	21121	0.8692
22222	0.7144	21121	0.7637	11112	0.8572	11221	0.8930
21121	0.7543	11211	0.7914	11222	0.8768	21111	0.9035
22332	0.7859	21211	0.8117	21122	0.8891	22222	0.9126

The most frequent 3L profile is 11111 ('full health') for the HSE and PROMs after surgery. All samples capture more than 90% of observations in the 10 most frequent profiles. HSE and after surgery Groin Hernia and Varicose Vein are the most concentrated, with 75% of observations in 3 profiles. 11111 and 11121 are the most frequent profiles in these 3 samples, reported by 70%.

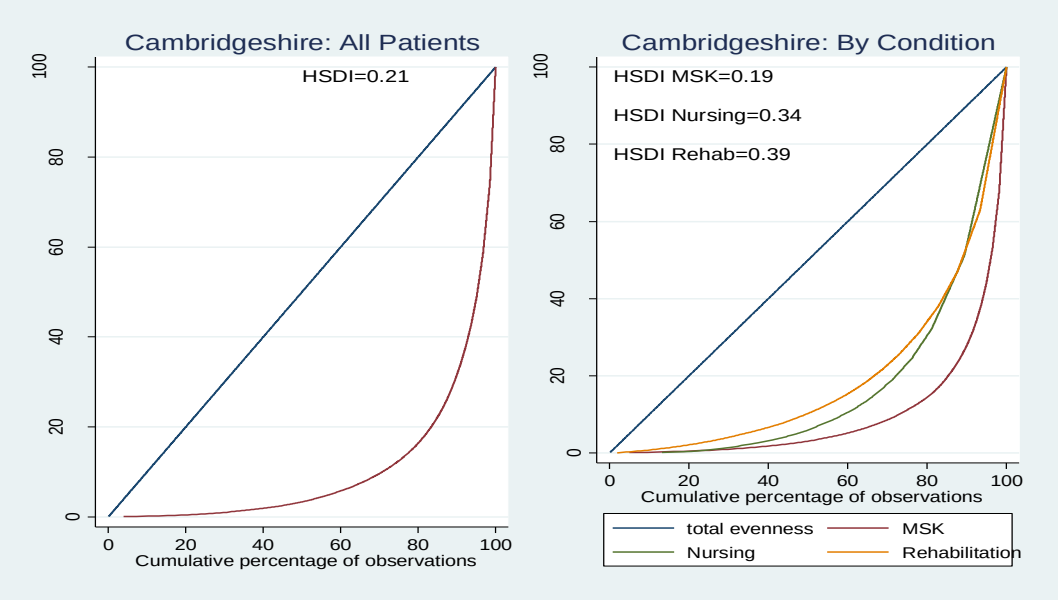
4.2. Applying the HSDI and Shannon indexes to EQ-5D profiles

As explained in section 3, two set of indexes are reported: the Shannon index H', which increases according to absolute informativity or absolute discriminatory power of the EQ-5D to classify patients in each dataset; and a second set of indexes: HSDI and Shannon evenness J', which decrease with the degree of clustering of the distribution, from 0 for maximum clustering/inequality to 1 for

minimum clustering/maximum equality. The J' index is also interpreted as relative discriminatory power or evenness index, which is maximum for total evenness or uniform distribution of patients across health states, in which case the value of J' is one and the value of H' is maximum at ln(S).

Figures 3, 4, and 5 show the HSDC. The HSDC shows that EQ-5D-5L profiles are most concentrated for MSK patients, and the HSDC crosses for Rehabilitation and Specialist Nursing although it shows less concentration for Rehabilitation in the most frequent profiles. Consequently, the larger HSDI is obtained for Rehabilitation patients.

Figure 3. HSDC: EQ-5D-5L profiles from NHS Cambridgeshire patients



The HSDC for the 3L is more concentrated than that of the 5L, with lower HSDIs. The HSDC for PROMs before and after surgery profiles almost overlap but the after data are more concentrated and have a lower HSDI.

Figure 4. HSDC: EQ-5D-3L profiles from HSE

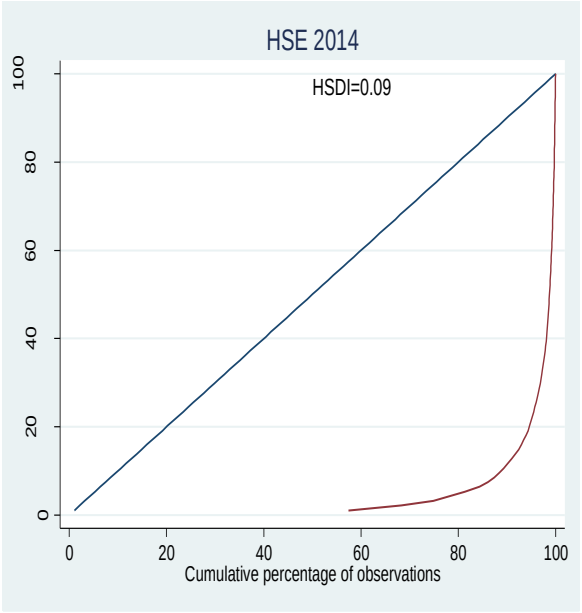


Figure 5. HSDC: EQ-5D-3L profiles from NHS PROMs

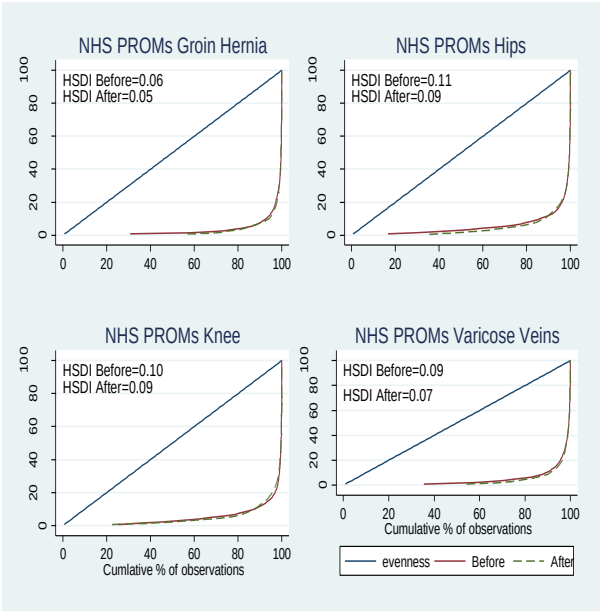


Table 8. HSDI and Shannon indexes from EQ-5D profiles

	Number of profiles S	Shannon index H'	HSDI	Shannon Evenness J'	Bias H'
EQ-5D-5L profiles					
Cambridgeshire NHS: All Patients	1,730	5.89	0.21	0.79	0.05
Cambridgeshire NHS: MSK	1,141	5.41	0.19	0.77	0.08
Cambridgeshire NHS: Specialist Nursing	732	5.38	0.34	0.82	0.46
Cambridgeshire NHS: Rehabilitation	1,240	6.39	0.39	0.90	0.23
EQ-5D-3L profiles					
Health Survey for England 2014	94	1.92	0.09	0.42	0.02
NHS PROMs Hernia – Before	143	2.30	0.06	0.46	0.00
NHS PROMs Hernia – After	140	1.89	0.05	0.38	0.00
NHS PROMs Hips – Before	147	3.00	0.11	0.60	0.00
NHS PROMs Hips – After	158	2.71	0.09	0.54	0.00
NHS PROMs Knee – Before	147	2.87	0.10	0.57	0.00
NHS PROMs Knee – After	153	2.85	0.09	0.57	0.00
NHS PROMs Varicose Veins - Before	104	2.32	0.09	0.50	0.01
NHS PROMs Varicose Veins - After	98	1.89	0.07	0.41	0.01

The estimated H' values confirm that the absolute discriminatory power of the EQ-5D-5L is larger than that of the 3L (see Table 7). This is a direct consequence of the larger number of observed profiles in the 5L, but may also reflect a more even distribution, implying better discrimination between patients. This is confirmed by comparing the 5L and 3L results for the evenness indexes J'

and HSDI (see Table 7). Both indexes are lower for the 3L. Moreover, the values of the J' index are closer to 1 than those of HSDI confirming the different sensitivity of the indexes to the proportion of rare profiles, with J' index being less affected by the addition of rare profiles which are the majority of profiles for the EQ-5D, as reported by only one or two patients. On the contrary, the HSDI decreases steeply capturing the inequality added by rare profiles.

With regard to discriminatory power, the largest H' is achieved from NHS Cambridgeshire Rehabilitation patients, with H' = 6.39; adjusting for underestimation bias, the true H for this group could be 6.6, close to the maximum H' for the uniform distribution (7.1 = ln(1240)). The relative discriminatory power of the profile is very large, reflected by J' = 0.9.

To compare the discriminatory power between NHS Cambridgeshire MSK and Specialist Nursing, we must consider the larger H' bias of Specialist Nursing. Taking this into account, the EQ-5D-5L discriminates better both in absolute and in relative terms for Specialist Nursing patients than for MSK patients, according to all 3 indexes. Of note is that the difference in the HSDI between MSK and Specialists Nursing is greater than that of J'. We analyse in the next subsection the different effects of rare profiles in both groups.

For the 3L profiles, there is a general pattern of greater concentration of the healthiest profiles (HSE, Groin Hernia after surgery, and Varicose Vein after surgery) related to less discriminatory power as reported by the three indexes H', J', and HSDI. The exception is that the HSDI for the HSE is larger than for most of the PROMs profiles, because it contains fewer rare profiles; a total of S=94 different profiles compared to between 98 and 158.

Health improvement after surgery is associated with decreased absolute and relative discriminatory power, except that J' for Knee surgery does not change.

4.3. Behaviour of HSDI and J' indexes to concentration and rareness

We analyse here whether the linear relationship between relative indexes (HSDI and J') and the proportion of rare profiles we found in the simulated profiles also holds in the population and patient data. Therefore, the linear regression parameters would inform us on the degree of concentration and on whether the differences between the indexes can be attributed to random variations in the number of patients versus rare profiles. The results are presented in Table 8.1 and Table 8.2.

Table 8.1 Linear regression of HSDI and J' on proportion of rare profiles

	EQ-5D-5L Cambridgeshire data sets								EQ-5D-3L dataste	
	All patients		MSK		Specialist Nursing		Rehabilitation		HSE 2014	
	HSDI	J	HSDI	J	HSDI	J	HSDI	J	HSDI	J
slope	-0.7460	-0.2581	-0.7564	-0.2827	-0.7229	-0.0861	-0.8862	-0.1856	-0.8306	-0.4485
intercept	0.8053	1.0046	0.8024	1.0072	0.8466	0.8828	0.8356	0.9959	0.8596	0.8560
R2	0.9993	0.9782	0.9921	0.9796	0.9749	**0.2900**	0.9886	0.9941	0.9817	0.8910
S	1730	1730	1141	1141	732	732	1240	1240	94	94
Index (at s=2, s=S)	(0.95, 0.21)	(0.99, 0.79)	(0.98, 0.19)	(0.99, 0.77)	(0.73, 0.34)	**(0.78, 0.82)**	(0.89, 0.39)	(0.96, 0.90)	(0.66, 0.09)	(0.63, 0.42)
proportion rare (at s=S)		0.791		0.803		0.703		0.522		0.9276
(min.- max) obs per profile		(1-1201)		(1-990)		(1-444)		(1-129)		(1-4069)
N (total obs.)		30,284		19,999		3,366		6,919		7,085

Table 8.2 Linear regression of HSDI and J' on proportion of rare profiles

	EQ-5D-3L datasets							
	PROMS Hernia Before		PROMs Hernia After		PROMs Hips Before		PROMs Hips After	
	HSDI	J	HSDI	J	HSDI	J	HSDI	J
slope	-0.8382	-0.5546	-0.8543	-0.5188	-0.8471	-0.4525	-0.8352	-0.4512
intercept	0.8410	1.0042	0.8739	0.9031	0.8241	1.0139	0.8466	0.9724
R2	0.9827	0.9770	0.9864	0.8990	0.9854	0.9853	0.9918	0.9204
S	143	143	140	140	147	147	158	158
Index (at s=2, s=S)	(0.97,0.06)	(0.99,0.46)	(0.68,0.05)	(0.69,0.38)	(0.92,0.1)	(0.98,0.6)	(0.72,0.09)	(0.76,0.54)
proportion rare (at s=S)		0.9303		0.9528		0.8634		0.9047
(min.- max) obs per profile		(1-14539)		(1-26831)		(1-11507)		(1-24157)
N (total obs.)		47,276		47,168		69,132		67,924
	EQ-5D-3L datasets							
	PROMS Knee Before		PROMs Knee After		PROMs VV Before		PROMs VV After	
	HSDI	J	HSDI	J	HSDI	J	HSDI	J
slope	-0.8361	-0.4604	-0.8230	-0.4732	-0.8291	-0.5095	-0.8399	-0.5053
intercept	0.8247	1.0048	0.8186	1.0100	0.8327	0.9803	0.8579	0.9052
R2	0.9912	0.9672	0.9864	0.9795	0.9835	0.9798	0.9847	0.9390
S	147	147	153	153	104	104	98	98
Index (at s=2, s=S)	(0.79,0.1)	(0.87,0.57)	(0.91,0.09)	(0.97,0.57)	(0.89,0.09)	(0.96,0.5)	(0.73,0.07)	(0.77,0.41)
proportion rare (at s=S)		0.8802		0.8871		0.9024		0.9324
(min.- max) obs per profile		(1-18281)		(1-15962)		(1-6247)		(1-9549)
N (total obs.)		72,423		71,084		17,658		17,597

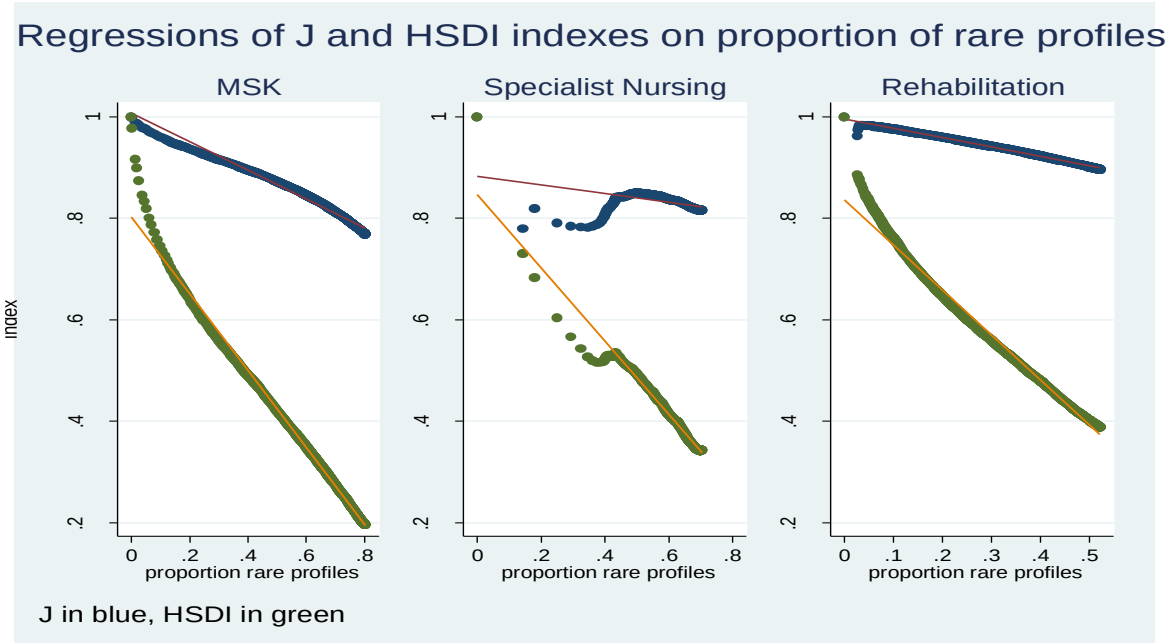
The regressions presented in Table 8.1 and Table 8.2 confirm the pattern found in Table 2 from the simulated samples with an important exception which we analyse in more detail; this exception occurs in the sample of Specialist Nursing patients reporting the EQ-5D-5L profiles in Cambridgeshire health centres.

Apart from the case of Cambridgeshire Specialist Nursing, all regressions present almost perfect goodness of fit, with the lowest R-squared=0.89 found for the regression of J' from HSE 2014. Similar values of the R-squared below 1 are found for the regressions of J' in PROMs Hernia and Varicose Vein after surgery. These three distributions of the EQ-5D-3L are characterised by a large ceiling with more than half of sample observations (see Table 6) which can affect a larger decrease of J in the first few profiles. This is also shown by the intercept estimate being slightly smaller than one in the J' regressions (around 0.90).

Again, HSDI is more sensitive to rare profiles than J' index and it shows an almost constant slope in the interval (-0.8, -0.9), so that the proportion of rare profiles observed around 0.9 produces a decrease of 0.72 in the HSDI, rendering the index in its final values around 0.1 for the 3L samples. The largest values of the HSDI in the 5L samples are mainly explained by a smaller importance of rare profiles, especially in Rehabilitation patients with only 52% of rare profiles. The values of J' around 0.4 found in the EQ-5D-3L samples are also explained by the larger sensitivity of J' to rare profiles, with negative slope around -0.5. However, in the 5L samples, J' is larger due to both a lower proportion of rare profiles and a flatter pattern of the linear regression.

The case of Cambridgeshire Specialist Nursing sample is remarkable and cannot be detected by merely comparing the indexes, which are similar to those obtained from Rehabilitation patients. However, the regression pattern of J' presents a fundamental difference with a very low goodness of fit (R-squared=0.29) hinting a nonlinear relation between the index and the proportion of rare profiles. This nonlinear relation can be observed in Figure 6 which presents the data and fitted regressions for the three Cambridgeshire groups.

Figure 6. Relationship between the indexes and proportion of rare profiles



The case of Specialist Nursing can be described as a sample with two groups of patients. A first group of around one-fourth of the sample is composed of a concentrated distribution of just 5 profiles. And the rest of the patients are almost uniformly distributed across the rest of the observed profiles (727 profiles). Both the J' and HSDI index decrease when considering patients in the first few profiles, up

to s=5, with a high concentration of the first three profiles (with 444, 133, and 115 patients respectively). However, between the profile s=6 and the final profile s=732, the distribution of patients approaches a uniform distribution of 2,569 patients distributed across 727 profiles, with a maximum of 44 patients, and a minimum of 1 patient per profile. This explains the increase in both indexes (J' after the 5th profile, and HSDI after the 10th profile), which is maintained for the J' index given that it is not sensitive to random variations in the number of patients.

4.4. Power Law Analysis

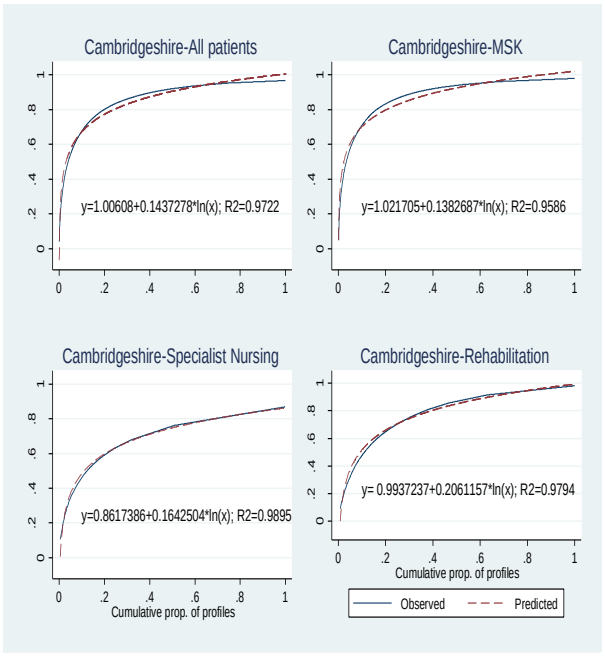
The power law is estimated as in regression (11) where the dependent variable is the cumulative proportion of observations and the explanatory variable is the log of cumulative proportion of profiles. The estimates for the different data sets using Ordinary Least Squares are shown in Figures 7 to 9, calculated excluding the profile 11111 for both 5L and 3L profiles and rare profiles with one or two patients for the 3L. As shown in the estimations from simulated samples (Table 3 and Figure 2), this allows to reduce the proportion of rare observations and improve the goodness of fit of the power law. The goodness of fit as measured by R-squared is always good; it is better for the 5L than the 3L data.

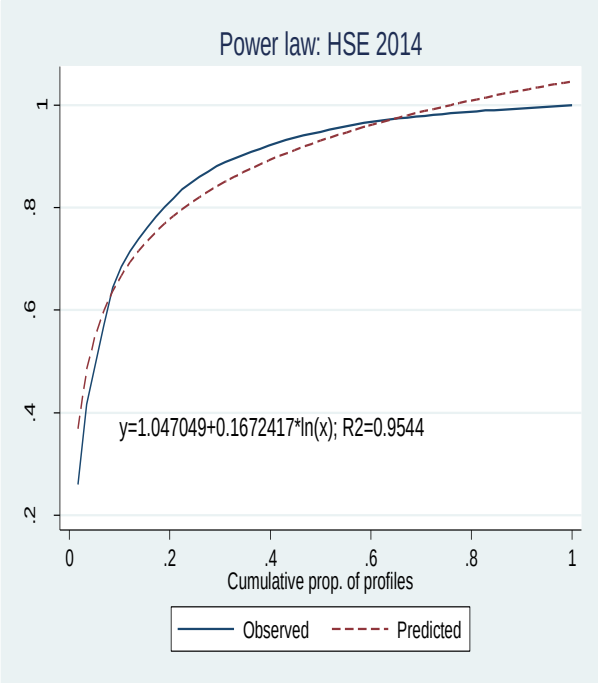
The constant coefficient *a* is correctly estimated at around 1, reflecting the fact that the total set of profiles must contain all observations (x = 1; log(1)=0; y=1). The estimated values for *y* using OLS can give impossible values >1 at higher levels of *x*, showing a limitation of this simple estimator.

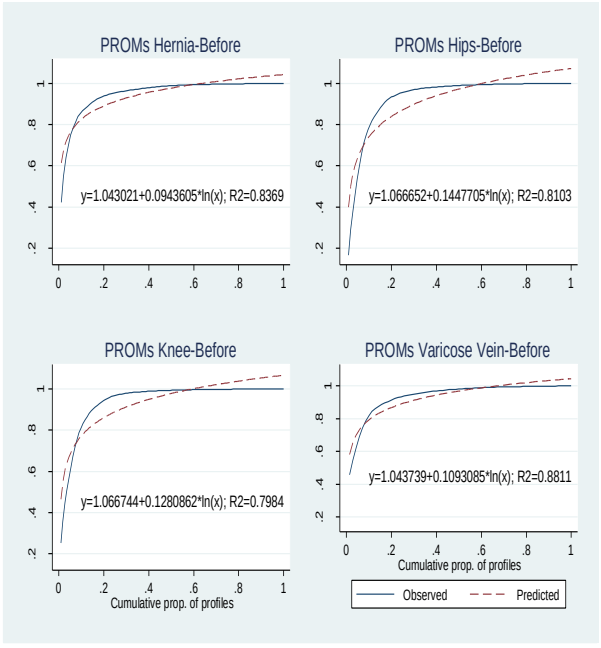
The *b* coefficient shows the sensitivity of the power law relationship to the concentration of profiles. Because the cumulative proportion of profiles lies between 0 and 1, its logarithm is negative. So, the larger the slope *b* is, the less concentrated is the distribution; any cumulative proportion of profiles predicts a smaller cumulative proportion of observations.

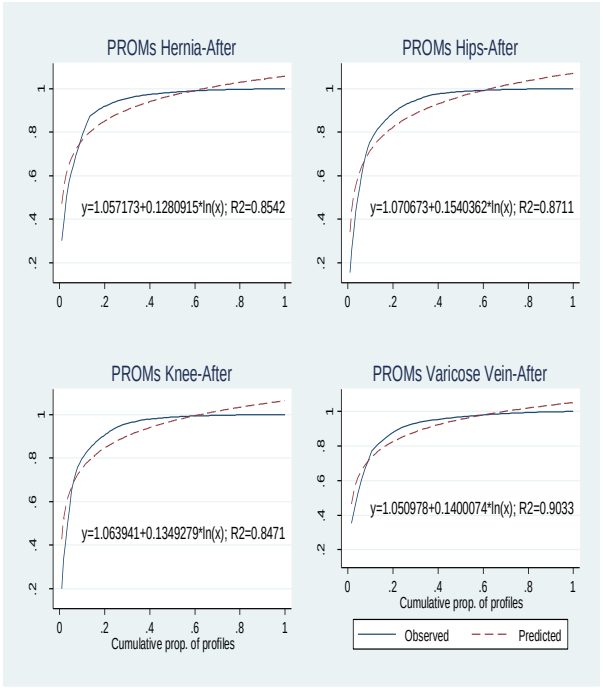
For the EQ-5D-5L, the largest *b* coefficient is for rehabilitation, suggesting that this is the least concentrated distribution, which is consistent with the results presented from the other methods. For the EQ-5D-3L, the largest *b* coefficients are for Knee and Hip surgery data. They are larger after than before surgery but this result is strongly affected by the removal of the ceiling 11111 given that a remarkable effect of the surgery is evidenced by the increase in the ceiling (see Table 6).

Figures 3-9: Power Law Analysis









5. DISCUSSION AND CONCLUSIONS

In this paper, we have reported new methods for describing the distribution of EQ-5D profiles diagrammatically and via summary indices and power functions. Together with other indices available in the literature, these provide a set of techniques which complement existing methods for describing EQ-5D profile data. Our exploration of the properties of these indices show that they each perform somewhat differently with respect to their sensitivity to particular aspects of observed distributions.

Our hope is that these methods will encourage those collecting EQ-5D data to pay more attention to the profile data provided by patients. While we have focussed our exploration of these methods on EQ-5D data, the methods are equally applicable to other PRO instruments.

Our methods may also have applications in describing the concentration or dispersion of self-reported health in patient groups, in a manner than can inform their clinical management. For example, if EQ-5D data are highly concentrated on a few profiles, this suggests that patient management can make effective use of standardised treatment protocols. In contrast, where patients' profile data are dispersed over a large number of profiles, this suggests individualised approaches to treatment may be required. The use and comparison of both indexes, HSDI and J', can be used to assess this degree of concentration. For example, imagine a situation where there are two different patient groups and reported measures for HSDI and J' for each group. In the first group both indexes J' and HSDI are close to zero but in the second group only HSDI is close to zero while J' is larger or closer to one. Since J' is less sensitive to random variation in the number of patients per profile, the conclusion is that changes in the number of patients per profile reflect better the prevalence of patient problems in the first group than in the second group. The conclusion is that there are more rare patients in the first than in the second group and these would benefit from personalised management, while standardised treatment protocols could be applied to the types of patients in the second group.

Our application of these methods to NHS data provides an additional way of investigating the discriminatory power of the EQ-5D-5L compared to the EQ-5D-3L in both general and patient populations. Our methods can describe how the EQ-5D instruments differentiate between patients in comparisons within or across diseases, or before and after treatment. The results reported here supports the conclusion reported elsewhere (Janssen et al., 2013; Janssen, Bonsel and Luo, 2018) that the added levels of the EQ-5D-5L allows better differentiation between patients. However, an additional observation of the analyses reported in this paper is that this occurs within the most frequent profiles rather than by generating more rare profiles. In fact, the importance of rare profiles is lower in the datasets we explore for the 5L than in those of the 3L version. Also, the

relative importance of rare profiles increases with the level of health gained after the four interventions considered in PROMs data. This is associated with more highly concentrated PROMs EQ-5D-3L profile data after than before surgery, as patients shift from a (wider) range of poor health states to converge on (narrower) set of milder states.

These same approaches could also be used to compare the nature of the health state descriptions generated from other PRO instruments. The methods described in this paper are best thought of as a complement to existing methods for analysing PRO data, and provide a way of gaining additional insights into the nature of PRO data obtained from patients. The methods have their own limitations – for example, the specific J' or HSDI calculated for a sample will be driven purely by the frequency with which profiles are observed, but is indifferent to other aspects of those profiles. For example, two samples could each have the same HSDI (or J') yet in one, the most common profile be a severe health state, and in the other a very mild health state. Simple descriptive approaches to describing profile data therefore continue to be important. The principal contribution of this paper is to offer a set of methods to capture the degree of clustering and dispersion of observed health states, a property of data sets that, to date, has not adequately been reported or analysed.

REFERENCES

Agborsangaya, C. B., Lahtinen, M., Cooke, T. and Johnson, J. A., 2014. Comparing the EQ-5D 3L and 5L: measurement properties and association with chronic conditions and multimorbidity in the general population. *Health and quality of life outcomes,* 12 (1), 1.

Ball, P., 2005. Critical Mass: How one thing leads to another. London: Arrow Books.

Devlin, N., Parkin, D. and Browne, J., 2010. Patient-reported outcome measures in the NHS: new methods for analysing and reporting EQ-5D data. *Health Economics,* 19 (8), 886-905.

Dolan, P., 1997. Modeling valuations for EuroQol health states. *Medical care,* 35 (11), 1095-1108.

Feng, Y., Devlin, N., Bateman, A., Zamora, B. and Parkin, D., 2016. The Distribution of the EQ-5D-5L Index in Patient Populations. *OHE Research Paper*.

Feng, Y., Parkin, D. and Devlin, N. J., 2014. Assessing the performance of the EQ-VAS in the NHS PROMs programme. *Quality of Life Research,* 23 (3), 977-989.

Gini, C., 1921. Measurement of inequality of incomes. *The Economic Journal,* 31 (121), 124-126.

Gutacker, N., Bojke, C., Daidone, S., Devlin, N. and Street, A., 2013. Hospital variation in patient-reported outcomes at the level of EQ-5D dimensions: evidence from England. *Medical Decision Making,* 33 (6), 804-818.

Janssen, M., Pickard, A. S., Golicki, D., Gudex, C., Niewada, M., Scalone, L., Swinburn, P. and Busschbach, J., 2013. Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. *Quality of Life Research,* 22 (7), 1717-1727.

Janssen, M. F., Bonsel, G. J. and Luo, N., 2018. Is EQ-5D-5L better than EQ-5D-3L? A head-to-head comparison of descriptive systems and value sets from seven countries. *PharmacoEconomics,* 1-23.

Jost, L., 2007. Partitioning diversity into independent alpha and beta components. *Ecology,* 88 (10), 2427-2439.

Lande, R., 1996. Statistics and partitioning of species diversity, and similarity among multiple communities. *Oikos,* 5-13.

Lorenz, M. O., 1905. Methods of measuring the concentration of wealth. *Publications of the American statistical association,* 9 (70), 209-219.

Molinari, J., 1989. A calibrated index for the measurement of evenness. *Oikos,* 319-326.

Parkin, D., Devlin, N. and Feng, Y., 2016. What Determines the Shape of an EQ-5D Index Distribution? *Medical Decision Making,* 0272989X16645581.

Parkin, D., Rice, N. and Devlin, N., 2010. Statistical analysis of EQ-5D profiles: does the use of value sets bias inference? *Medical Decision Making,* 30 (5), 556-565.

Pielou, E., 1966. Shannon's formula as a measure of specific diversity: its use and misuse. *The American Naturalist,* 100 (914), 463-465.

Pielou, E., 1975. Ecological diversity. *J. Wiley and Sons, New York*.

Shannon, C. E., 1948. A mathematical theory of communication. *Bell System Technical Journal,* 27 (3), 379-423.

Smith, B. and Wilson, J. B., 1996. A consumer's guide to evenness indices. *Oikos,* 70-82.

Von Steinbuechel, N., Covic, A., Polinder, S., Kohlmann, T., Cepulyte, U., Poinstingl, H., Backhaus, J., Bakx, W., Bullinger, M. and Christensen, A.-L., 2016. Assessment of Health-Related Quality of Life after TBI: Comparison of a Disease-Specific (QOLIBRI) with a Generic (SF-36) Instrument. *Behavioural neurology,* 2016.

Whittaker, R. H., 1972. Evolution and measurement of species diversity. *Taxon,* 213-251.