The importance of normalisation in the construction of deprivation indices

Mark S Gilthorpe

Abstract

Study objectives – Measuring socio-economic deprivation is a major challenge usually addressed through the use of *composite indices*. This paper aims to clarify the technical details regarding composite index construction. The distribution of some variables, for example unemployment, varies over time, and these variations must be considered when composite indices are periodically re-evaluated. The process of normalisation is examined in detail and particular attention is paid to the importance of symmetry and skewness of the composite variable distributions.

Design – Four different solutions of the Townsend index of socioeconomic deprivation are compared to reveal the effects that differing transformation processes have on the meaning or interpretation of the final index values. Differences in the rank order and the relative separation between values are investigated.

Main results – Constituent variables which have been transformed to yield a more symmetric distribution provide indices that behave similarly, irrespective of the actual transformation methods adopted. Normalisation is seen to be of less importance than the removal of variable skewness. Furthermore, the degree of success of the transformation in removing skewness has a major effect in determining the variation between the individual electoral ward scores. Constituent variables undergoing no transformation produce an index that is distorted by the inherent variable skewness, and this index is not consistent between re-evaluations, either temporally or spatially.

Conclusions – Effective transformation of constituent variables should always be undertaken when generating a composite index. The most important aspect is the removal of variable skewness. There is no need for the transformed variables to be normally distributed, only symmetrically distributed, before standardisation. Even where additional parameter weights are to be applied, which significantly alter the final index, appropriate transformation procedures should be adopted for the purpose of consistency over time and between different geographical areas.

(J Epidemiol Comm Health 1995;49(Suppl 2):S45–S50)

The Department of Public Health and Epidemiology, University of Birmingham, Birmingham B15 2TT
M S Gilthorpe

Correspondence to:
Dr M S Gilthorpe.

In the fields of public health, epidemiology and the social sciences, the development and application of composite indices is becoming increasingly common. Abstract measures, in the form of scales or ratios, are invaluable in quantifying conditions that would otherwise require a more qualitative or descriptive approach. To some, this aspiration to denote complex conditions numerically is considered perverse; their concerns are that these abstractions lead to the index becoming the main focus with the real issues being side stepped.[1] However, to proceed with much research work in the aforementioned fields, many aspects of our society have been quantified in this manner.

Quantifying complex conditions or circumstances not only enables comparisons to be made, but also facilitates the development of theories and formulas. Examples of this are seen in the investigation of socioeconomic deprivation and its impact upon the health needs of a population.[2-13] Measuring deprivation is a major challenge to researchers, and the published reports record many varied approaches.[14 15] Nearly all resort to the creation of a composite index, although some argue for the use of a single variable such as unemployment.[16 17]

To measure any one social characteristic there are many variables that could be used. Constituent variables are deliberately chosen for their association to the subject matter under study, and often large intercorrelations exist between them. Deprivation indices are usually derived from a combination of small area statistics (SAS), extracted from OPCS census data.[18] The additive normal process allows the useful combination of these variables. The combination of different variables could yield a wide range of composite indices that might disagree vastly in absolute score. However, it is not the intention of this study to question the choice of constituent variables but to examine more closely the construction of any single index.

If the distribution of a constituent variable is skewed, the measure of its dispersion, the SD, is exaggerated. In such instances, applying standardisation will result in a distortion of the perceived relative differences across this variable when compared with others. Within the additive normal process, the crude variables are normalised to yield a more symmetric distribution before standardisation is applied. However, the various normalisation procedures applied to each variable could effect the final index distribution. When an index is to be used within formulas, the distribution or range of values is of greater importance. Often emphasis is placed on the relative difference, or the rank

order, of index values, rather than their absolute magnitude.

This paper investigates the process of normalisation in detail through examining a commonly used deprivation index that is generated using the additive normal process, namely the Townsend index of socioeconomic deprivation.[8] In particular, the effects upon the rank order and the relative differences across the index are examined.

Methods

THE TOWNSEND INDEX

The Townsend deprivation index consists of four population variables: percentage measures of local unemployment, car ownership, overcrowding, and home ownership. For the definition and construction of each census variable, see Appendix A. These values are derived from the 1991 census and are calculated for 826 electoral wards within the West Midlands Regional Health Authority (WMRHA). When first evaluating this index using the 1981 national census data, Townsend transformed two of the variables, those pertaining to unemployment and overcrowding, using the natural logarithm transformation $y = \ln(x+1)$ to produce more normal distributions.

NORMALISATION OF THE CONSTITUENT VARIABLES

The additive normal process applies a transformation to each variable to produce values that are more normally distributed. The transformed variables are standardised and combined to produce a single index. However, it is not a requirement of the deprivation index that census variables should be normally distributed. The purpose of the transformation is to achieve values that are symmetrically distributed, to allow standardisation to follow. Hence, the "normalisation process", in this context, is a loose term for what is perhaps more appropriately referred to as the "transformation process", which minimises variable skewness. Variable normality, per se, is therefore of no concern.

The transformation process requires the application of a continuous function to each of the composite variables. These functions may differ and the exact choice of transformation is arbitrary. Many of the variables chosen to form composite indices are strictly positive, and measure continuously from zero. Therefore, the crude variables tend to be positively skewed. For this reason, transformations that are renowned for reducing positive skewness should be used. The most commonly used transformations are the square root, reciprocal and logarithm; the power of correctional adjustment for skewness increasing correspondingly. To evaluate the success of any transformation, the resulting variable skewness is considered.[19] Ideally, the transformation should result in a smooth, symmetrical distribution, without large peaks, or clumps of values at the extremes. To assess the degree of smoothness of the transformed distribution, a measure of kurtosis is used.[19] For the definitions of skewness and kurtosis, see Appendix B.

When applying these transformations, care has to be taken over the occurrence of zero or negative values. In practice there is no restriction on the type or number of functions that may be considered, providing that each transformation is a continuous one to one map across the data interval. Without this constraint the nature and intrinsic qualities of the variable are prone to corruption or alteration. The square root function, when applied to positive values, is a one to one map provided that only positive or negative roots are chosen throughout; the logarithm cannot be used unless the data values are strictly positive; and the reciprocal can only be applied to non-zero data. If these conditions are not satisfied, an offset may be added to the variable to correct this. After transformation, variable intercorrelations may become negative. To alleviate this, all transformed values must be multiplied by a negative one.

As standardisation is a symmetrical process about the distribution mean, it is important to minimise the measure of skewness while maintaining the smallest possible measure of kurtosis. To achieve this, a spreadsheet program was written within the package *Quattro-Pro* to perform transformations on each of the four variables.[20] All Townsend statistics are percentages and are therefore positive; variable four, however, contains zeros. For this reason, an arbitrary initial offset of unity is used as default for all transformations applied to this variable. Several composite transformations are explored and these are listed in table 1, along with their resulting measures of skewness and kurtosis. Table 1 also presents the range of values that result from each transformation. For a zero offset, transformations are identical if consecutive square root and reciprocal functions are interchanged, since these functions are commutative.

For each of the three transformations, the offsets were altered slightly, and improvements in skewness and kurtosis were noted. In turn, each function is optimised by the adjustment of the offset parameter. This is achieved entirely by "trial and error", with attention also being paid to the range of transformed values. If the value range is extremely small, this solution is rejected and another is sought. This is to avoid having to apply standardisation to very small numbers. A trial may begin with an offset of say "one" and the resulting variation in skewness is observed. Those functions that have shown a large reduction in skewness are noted. Some transformations may result in a negatively skewed distribution. In such instances, an intelligent guess is made for an offset value between one and zero. The trials proceed until the desired transformation yields a distribution with zero skewness.

It is observed that for some transformations there is no ideal solution (that being zero skewness and minimal kurtosis), but for others there are several. For all transformations, however, the variation in skewness or kurtosis is erratic and could generally be described as chaotic.

Table 1 The composite functions used to transform the four census variables; illustrating variable skewness, kurtosis, and range. The transformations are combinations of the reciprocal (R), the natural logarithm (Ln), and the square root (S) functions

Function	Unemployment			Car ownership			Home ownership			Overcrowding		
	Skewness	Kurtosis	Range	Skewness	Kurtosis	Range	Skewness	Kurtosis	Range	Skewness	Kurtosis	Range
Crude	2·05	5·68	30·96	0·79	−0·14	69·13	0·60	0·11	69·19	3·07	15·52	13·51
R	0·69	1·33	0·48	1·30	3·18	0·24	4·22	28·48	0·38	0·53	0·20	0·93
Ln	0·59	0·11	2·83	−0·07	−0·90	2·93	−0·88	1·40	3·35	0·69	0·96	2·67
S	1·25	1·79	4·34	0·37	−0·82	6·57	−0·04	−0·10	6·88	1·64	5·00	2·81
RLn	−0·59	0·11	2·83	0·07	−0·90	2·93	0·88	1·40	3·35	0·22	−0·20	0·63
	1·21	4·67	1·21	0·95	1·75	0·50	3·84	26·78	0·85	0·67	0·77	0·73
RS	0·02	−0·09	0·54	0·57	−0·08	0·39	2·15	7·74	0·51	−0·49	0·38	0·29
SR	0·02	−0·09	0·54	0·57	−0·08	0·39	2·15	7·74	0·51	0·37	−0·04	0·38
SLn	0·27	−0·06	1·05	−0·23	−0·71	0·91	−1·30	3·09	1·11	0·34	0·30	0·92
LnS	0·59	0·11	1·41	−0·07	−0·90	1·47	−0·88	1·40	1·67	0·98	1·84	0·88
S-LnR	0·27	−0·06	1·05	−0·23	−0·23	0·91	−1·30	3·09	1·11	−0·48	0·18	0·41
LnRS	−0·59	0·11	1·41	0·07	−0·90	1·47	0·88	1·40	1·67	−0·60	0·63	0·22
LnSR	−0·59	0·11	1·41	0·07	−0·90	1·47	0·88	1·40	1·67	0·29	−0·14	0·17
RSLn	0·57	1·55	0·69	0·66	0·39	0·38	2·65	13·03	0·56	0·05	0·04	0·16
SRLn	0·57	1·55	0·69	0·66	0·39	0·38	2·65	13·03	0·56	0·58	0·61	0·29
RLnS	1·21	4·67	2·42	0·95	1·75	1·01	3·84	26·78	1·70	−0·49	0·43	0·20

Table 2 The four index solutions used in this study, along with variable skewness, kurtosis and range. The variables V1, V2, V3, and V4 are the OPCS percentage values derived for each electoral ward (see Appendix A)

	Unemployment	Car ownership	Home ownership	Overcrowding
Solution 1				
(Standard):	R{LnS(V1) + 1·12}	Ln{S(V2) + 0·8}	S{V3 + 1·2}	R{Ln{S(V4 + 1·0) + 0·18} + 1·0}
Skewness	0·00	0·00	0·00	0·00
Kurtosis	−0·01	−0·94	−0·13	−0·04
Range	0·34	1·22	6·61	0·44
Solution 2				
(Alternative):	RS{Ln(V1) + 1·23}	Ln{R(V2) + 1·5}	Ln{V3 + 26·0}	R{S{V4 + 1·0} − 0·03}
Skewness	0·00	0·00	0·00	0·00
Kurtosis	0·02	−0·96	−0·25	−0·10
Range	0·27	2·63	1·23	0·77
Solution 3				
(WMRID):	Ln{V1 + 1·0}	Ln{V2 + 1·0}	S(V3)	Ln{V4 + 1·0}
Skewness	0·73	−0·02	−0·04	0·69
Kurtosis	0·28	−0·94	−0·10	0·96
Range	2·44	2·72	6·88	2·67
Solution 4				
(Crude):	V1	V2	V3	V4
Skewness	2·05	0·79	0·60	3·07
Kurtosis	5·68	−0·14	0·11	15·52
Range	30·96	69·13	69·19	13·51

The functions will sometimes possess local minima and maxima, or attain enormous values, suggesting a tendency towards infinity. Nevertheless, it is not difficult to find several optimum solutions by "trial and error", and these are recorded to be considered later.

COMPARISON OF THE VARIOUS TRANSFORMATION PROCESSES

After many trials for different offsets, two optimal series of solutions are chosen where skewness is zero and kurtosis is minimised to less than one in magnitude for each constituent variable. The only considerations given to the final choice of transformations adopted were: firstly, that each solution should exhibit a range of alternative functions and offset values and, secondly, that each illustrates contrasting transformation methods in order to highlight the constructional differences between them.

Before this study, the West Midlands Regional Information Department (WMRID) created its own version of the Townsend index, adopting similar transformations to those originally used for the 1981 census variables (West Midlands Regional Information Department – personal communication). Their solution incorporates an additional logarithm transformation for the measure of car ownership, and the home ownership variable is transformed by means of the square root function.

The criteria used in the construction of this index is not recorded, but the additional transformations were included primarily to improve upon the normality of the variable distributions and to reduce variable skewness. However, zero skewness was not attained for any constituent variable. This solution is included as it is evaluated from regional census data and used locally. All comparisons made within this study are relative to the West Midlands Region only.

It is debated elsewhere whether the crude values should be "normalised" or not.[21][22] Although contrary to statistical convention, this idea is not dismissed and an index is included where the variables have undergone no transformation. The transformation processes of all four solutions are presented in table 2, along with their measures of skewness, kurtosis, and the transformed variable range. These values are standardised and summed in the usual way. Without loss of generality, the first index solution is taken as a reference standard against which all others are compared.

The various transformations result in different distributions of the final index. Although the mean Townsend score is zero for each solution, the range and magnitude of individual values vary. This is a problem when attempting to compare individual ward scores directly. This can be avoided by a process of "index scaling". The absolute values of ward scores are summed over each solution to produce an
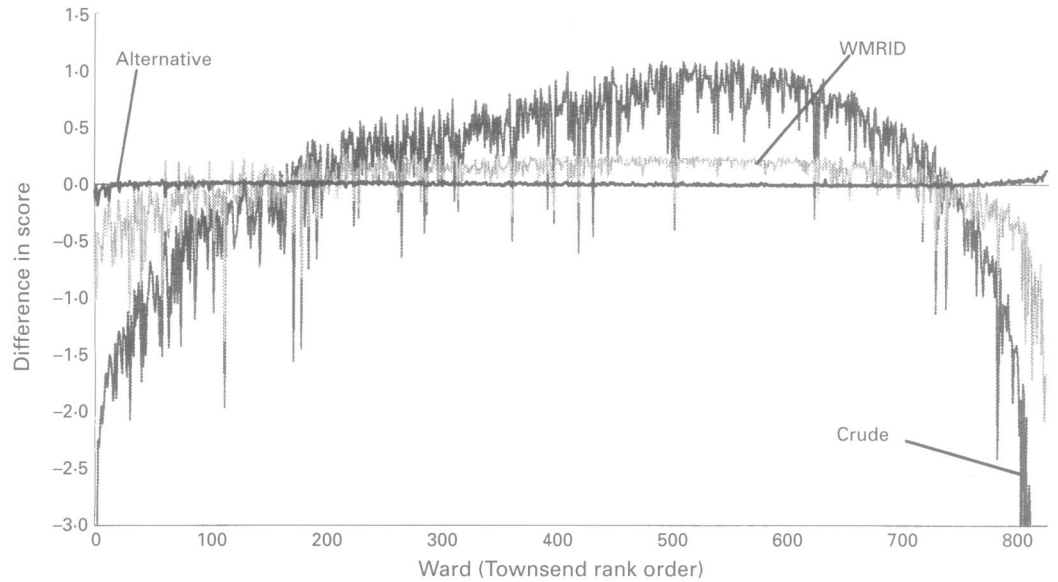
Figure 1   The index differences between the standard solution and other solutions, plotted in Townsend rank order of the standard solution. The crude solution is truncated; extreme values are −3·67 (left) and −10·86 (right).

"index weight". Using the standard solution as the reference, the remaining index solutions have their ward scores multiplied by the ratio of the standard index weight to that of their index weight. For example, for the alternative index solution this is illustrated algebraically:

$$\tilde{w}_a = \frac{\Sigma \, | w_s |}{\Sigma \, | w_a |} * w_a$$

where w_s is the standard ward score, w_a is the alternative ward score, $\tilde{w}_a$ is the adjusted or scaled alternative ward score, $|\;\;|$ denotes "absolute value", and summation is over all electoral wards. This process retains the rank position and relative range separation for each solution, while each index mean value remains at zero. The weighted values are ordered according to the rank order of the standard solution. Each ward score is subtracted from its corresponding standard value and the results are plotted in figure 1.
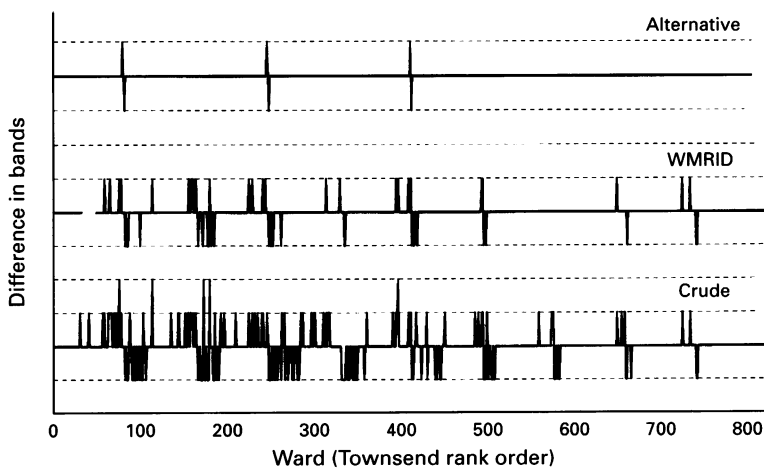

Figure 2   The index band differences between the standard solution and other solutions, plotted in Townsend rank order of the standard solution. The three comparisons are plotted on separate axes; each dotted line represents a deviation of one band.

In some studies, the composite indices are banded.[23–25] In these instances it is necessary to know how much the transformations can cause individual values to be placed into different bands. To examine this, the four solutions are aggregated into bands containing roughly equal numbers of wards. Ten bands of around 83 wards are chosen, corresponding to band widths ranging from 5·06, 4·94, 6·99, and 15·80 at the extremes of the distribution, to 0·88, 0·89, 0·84, and 0·80 around the mean Townsend value of zero, for the standard, alternative, WMRID, and crude solutions, respectively. The latter three solutions are differenced from the standard, and these differences are ordered according to the standard index rank order and plotted in figure 2.

Results

The two solutions chosen to satisfy the criteria of zero skewness and minimal kurtosis have undergone substantially different transformations. Despite this, it is clear from figure 1 that there is very little difference in the resulting indices. Further investigation of alternative solutions which meet the same criteria (not presented), again employing transformations that are different from those already illustrated, are consistent with this interpretation.

The WMRID index varies slightly from the standard solution; most noticeably at the extremes of the index range. The crude solution, however, is in very poor agreement with the standard solution and deviation is most pronounced at the outermost ranges of the index values. The extreme values of the crude index variation are truncated in figure 1. To the left of the range, the crude solution differs by −3·67; and to the right, the maximum difference from the standard solution is −10·86.

Another view of the variation across the indices is provided by grouping the index values

and plotting the subsequent band differences relative to the standard solution, as shown in figure 2. Clearly, some wards are being placed in different bands depending upon the transformation procedures adopted in generating the index. It is particularly important to notice the number of peaks and troughs across each plot. The alternative solution (variable skewness fully removed) places six wards into neighbouring bands from those of the standard solution, which results in disagreement in rank order on only three occasions. The other solutions illustrate more numerous differences and there is considerable disagreement in the rank order and the relative position of many ward values, across the entire range of Townsend values.

Discussion

Standardisation is generally acknowledged as an equitable process for combining several variables. Without it, disproportionate scales and ranges will give undue prominence to some variables at the expense of others. Even when the composite variables are uniformly expressed as percentages, the overall magnitude and range of each can differ greatly. Although usually resolved through "normalisation" and standardisation, this process has been challenged.[22] It has been stated that there is no inference from the index that an approximately normal distribution is required, and therefore it would be unreasonable to alter the original variables in this manner. However, it is the process of standardisation, and not the distribution of the final index, that vindicates the necessity of transformed constituent variables. Even after applying the additive normal process, the standard solution index has a skewness of 0·33 and kurtosis of −0·40.

The issue hinges on the desire to avoid giving undue prominence to some variables at the expense of others. The argument against transforming all variables to give equal weight to each is that the unweighted combination of the chosen parameters generates a better measure of that originally sought. If the property of skewness within a variable is to be deliberately exploited, the amalgamation of this variable with others may be better achieved without transformation (or without standardisation). However, this approach shows no methodological consistency, either temporally or spatially. It might be preferable to explore the properties associated with a skewed variable through its sole use, and not in composite form.[16]

The inclusion of parameter weights, as used by Jarman, yields a specially tailored index for which standardisation and therefore the removal of variable skewness are a prerequisite.[3] This is because, irrespective of the rationale behind their derivation, the Jarman weights must be applied to a set of variables that would have *initially* yielded an equal contribution to the final index. The additive normal process is therefore an appropriate method by which the Jarman index is generated, providing that sufficient care is taken over the removal of skewness.

Conclusions

It is necessary to consider carefully the methodologies employed when creating a new composite index, or re-evaluating an existing one using new data. When the additive normal process is used, issues concerning the transformation of composite variables are paramount. The standard and alternative solutions illustrate that, whatever transformations are employed, the removal of variable skewness leads to a consistent index (the differences between the standard and alternative solutions are very small). However, if skewness is not thoroughly removed, relative absolute differences between scores may occur (fig 1) even where consistency in rank order is maintained (fig 2). It is the distribution skewness, and not normality, that remains at the centre of the variable transformation debate.

Evaluation of a composite index may potentially introduce information that is purely an artefact of the chosen variable transformations. Careless choice of variable transformations can effectively colour the index. This has considerable implications upon the application of any formula that uses the absolute index values or their relative differences. In instances where a simple measure of socioeconomic deprivation is needed the crude Townsend solution may suffice, thus avoiding the need for a protracted transformation procedure. However, when an index informs policy decisions (or is used within formulae), a more precise algorithm must be adopted. This is particularly relevant where importance is placed upon individual ward scores, as these may vary considerably, depending upon the transformation processes adopted.

Appendix A

The 17 census statistics extracted from the Manchester database are: S080012, S080021, S080078, S080087, S080166, S080174, S080175, S080232, S080240, S080241, S210002, S210003, S230001, S230003, S230004, S230009, and S230013.[18] For each ward, these parameters generate four percentage variables, namely:

● V1 = % of economically active residents aged 16–59/64 (the retirement ages for men and women are 60 and 64 years respectively) who are unemployed

$$= 100 \times \frac{(S080078 - S080087) + (S080232 - S080240 - S080241)}{(S080012 - S080021) + (S080166 - S080174 - S080175)};$$

● V2 = % of private households who do not possess a car

$$= 100 \times \frac{S210003}{S210002};$$

● V3 = % of private households not owner occupied

$$= 100 \times \frac{S230001 - S230009 - S230013}{S230001};$$

● V4 = % of private households with more than one person per room

$$= 100 \times \frac{S230003 + S230004}{S230001}.$$

The four constituent variables are normalised to produce variables N1 ... N4, respectively. These are standardised, for example S1 = (N1-"*mean* of N1")/("*SD* of N1"). The composite index score is produced through the addition of these parameters: $T_{index} = S1 + S2 + S3 + S4$. For the purposes of this study, the index is standardised for WMRHA; hence these Townsend scores show deprivation relative to all wards in the West Midlands Region.

Appendix B

SKEWNESS

A measure of skewness (*s*) characterises the degree of asymmetry of a distribution around its mean value, and is given by:

$$s = \frac{n}{(n-1)(n-2)} \sum_{i}^{n} \left(\frac{x_i - \bar{x}}{SD} \right)^3,$$

where *n* is the number of distribution values (*n* must be >2); x_i (*i*=1,..,*n*) are individual distribution values; $\bar{x}$ is the mean value; and *SD* is the distribution standard deviation, which must not equal zero. A distribution with zero skewness is symmetric about the mean.

KURTOSIS

The kurtosis (*k*) of a data set measures a distribution's relative peakedness or flatness, indicating "heavy-tails" or clumped distributions. A kurtosis greater than zero is referred to as leptokurtic; a kurtosis less than zero is referred to as platykurtic. The kurtosis is given by:

$$k = \left\{ \frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum_{i}^{n} \left(\frac{x_i - \bar{x}}{SD} \right)^4 \right\} - \frac{3(n-1)^2}{(n-2)(n-3)},$$

where *n* is the number of distribution values (*n* must be greater than three); x_i (*i*=1,..,*n*) are individual distribution values; $\bar{x}$ is the mean value; and *SD* is the distribution standard deviation, which must not equal zero. A distribution with zero kurtosis is normally distributed.

1 Carr-Hill RA, Sheldon TA, Thunhurst C. The politics of deprivation indices and health. In: Spencer N, James H, eds. *Uses and abuses of deprivation indices.* Warwick: University of Warwick, 1992:24–32.

2 Carstairs V. Small area analysis and health services research. *Community Med* 1981;3:131–9.

3 Jarman B. Identification of underprivileged areas. *BMJ* 1983;286:1705–9.

4 Morgan M, Chinn S. ACORN group, social class and child health. *J Epidemiol Community Health* 1983;37:196–203.

5 Jarman B. Underprivileged areas: validation and distribution of scores. *BMJ* 1984;298:1585–8.

6 Thunhurst C. The analysis of small area statistics and planning for health. *The Statistician* 1985;34:93–106.

7 Balarajan R, Yuen P, Machin D. Socioeconomic differentials in the uptake of medical care in Great Britain. *J Epidemiol Community Health* 1987;41:196–9.

8 Townsend P, Phillimore P, Beattie A. *Health and deprivation: inequality in the north.* London: Routledge, 1988.

9 Curtis SE. Use of survey data and small area statistics to assess the link between individual morbidity and neigh-

bourhood deprivation. *J Epidemiol Community Health* 1990;44:62–8.

10 Ben-Shlomo Y, White I, McKeigue PM. Prediction of general practice workload from census based social deprivation scores. *J Epidemiol Community Health* 1992;46:532–6.

11 Eames M, Ben-Shlomo Y, Marmot MG. Social deprivation and premature mortality: regional comparison across England. *BMJ* 1993;307:1097–102.

12 Kelly WF, Mahmood R, Kelly MJ, Turner S, Elliot K. Influence of social deprivation on illness in diabetic patients. *BMJ* 1993;307:1115–6.

13 Reading R, Raybould S, Jarvis S. Deprivation, low birth weight, and children's height: a comparison between rural and urban areas. *BMJ* 1993;307:1458–62.

14 Morris R, Carstairs V. Which deprivation? A comparison of selected deprivation indexes. *J Public Health Med* 1991;13(4):318–26.

15 Dye TD, Lee RV. Socioeconomic status: developing a quantitative community based index in rural Kashmir. *J Epidemiol Community Health* 1994;48:421–2.

16 Campbell DA, Radford JMC, Burton P. Unemployment rates, an alternative to the Jarman index? *BMJ* 1991;303:750–5.

17 Payne JN, Coy J, Milner PC, Patterson S. Are deprivation indicators a proxy for morbidity? A comparison of the prevalence of arthritis, depression, dyspepsia, obesity and respiratory symptoms with unemployment rates and Jarman scores. *J Public Health Med* 1993;15(2):161–70.

18 Office of Population Census and Survey. *The 1991 census user's guide.* London: HMSO, 1993.

19 Rice JA. *Mathematical statistics and data analysis.* London: Wadsworth and Brooks, 1988:299–301.

20 Borland International Inc. *Quattro-Pro for Windows: version 5. User guide.* Scotts Valley, CA: Borland International Inc, 1993.

21 Carr-Hill RA, Sheldon TA. Designing a deprivation payment for general practitioners: the UPA(8) wonderland. *BMJ* 1991;302:393–6.

22 Martin D, Senior ML, Williams HCWL. On measures of deprivation and the spatial allocation of resources for primary health care. *Environment and Planning A* 1994;26:1911–29.

23 Billingham C, Thomas A, Gilman EA, Somervaille LJ, Griffiths RK. Resource allocation in the West Midlands: analysis at the ED level and the urban/rural split. In: Gilman EA, Munday S, Somervaille LJ, Strachan R, eds. *Resource allocation and health needs: from research to policy.* London: HMSO, 1994:86–90.

24 Gilman EA, Somervaille LJ, Griffiths RK. Deprivation and resource allocation in the West Midlands Regional Health Authority. In: Gilman EA, Munday S, Somervaille LJ, Strachan R, eds. *Resource allocation and health needs: from research to policy.* London: HMSO, 1994:19–25.

25 McLoone P, Boddy FA. Deprivation and mortality in Scotland, 1981 and 1991. *BMJ* 1994;309:1465–70.

Open discussion

GORDON – In the Townsend index some of the variables are more important as measures of deprivation than others. By adjusting down the skewness you can remove that difference. For the Townsend index, I think Townsend would argue that the variables that are skewed – unemployment, home ownership, and "no car" – are probably better indices than overcrowding, for example.

GILTHORPE – If that was the criterion by which he generated his index, then, I suggest, you would not transform those variables. It is arguably more important not to attempt to "normalise" the variables than to do it inappropriately.

DOLK – Whether you normalise the variables or not will affect considerably what happens to a particular electoral ward if you are going to allocate resources, because it might be misplaced in its index. But have you looked at the general relationship between deprivation indices and some health index, for example mortality, to see how much transformation affects that general relationship?

GILTHORPE – Not directly but I think you will find that the "noise" within the evaluation of the health index will yield similar disagreement in rank order when compared with the various deprivation index solutions.