Int J Public Health (2010) 55:71–74
DOI 10.1007/s00038-009-0078-y

HINTS & KINKS

Analysis of socioeconomic health inequalities using the concentration index

Peter Konings · Sam Harper · John Lynch · Ahmad Reza Hosseinpoor · Dirk Berkvens · Vincent Lorant · Andrea Geckova · Niko Speybroeck

Received: 28 May 2009 / Revised: 11 September 2009 / Accepted: 14 September 2009 / Published online: 2 October 2009

Although "average" health conditions at national and global levels may have improved, inequalities in health conditions still exist among and within countries (Mackenbach et al. 2008; Smits and Monden 2009). Epidemiologists often focus on average health and try to understand its determinants. This is important, but measuring and understanding inequalities in health, i.e. the gap in health between disadvantaged and better-off groups, is an additional concern for health policy makers, health communicators and health scientists.

P. Konings
Department of Electrical Engineering, SCD (SISTA), K.U. Leuven, Leuven, Belgium

S. Harper
Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, Canada

J. Lynch
Sansom Institute, University of South Australia, Adelaide, Australia

J. Lynch
Department of Social Medicine, University of Bristol, Bristol, UK

A. R. Hosseinpoor
World Health Organization, Geneva, Switzerland

D. Berkvens · N. Speybroeck (✉)
Institute of Tropical Medicine, Antwerp, Belgium
e-mail: nspeybroeck@itg.be

V. Lorant · N. Speybroeck
Institute of Health and Society, Université Catholique de Louvain, Brussels, Belgium

A. Geckova
Kosice Institute for Society and Health, Kosice, Slovak Republic

This Hints & Kinks paper describes the computation of the concentration index. This index not only provides an indicator of health inequality but can also be decomposed proportionally into contributions of different inequality of health determinants. In a subsequent issue of this journal, the technique of decomposing inequalities of health will be clarified.

The relative concentration index (RCI) summarizes relative inequality across the entire socioeconomic distribution rather than simply comparing extremes [an absolute concentration index (ACI) may also be calculated by multiplying RCI by the mean level of health (Wagstaff et al. 1991)]. The RCI belongs to a general class of indicators sometimes called measures of 'disproportionality' (Harper and Lynch 2007). Such measures all express inequality as a function of differences between shares of some health outcome compared with shares of the population. The RCI measures this disproportionality across ordered social groups, such as income or social class, and, therefore, reflects the direction of the social gradient in disease (Wagstaff et al. 1991).

Assume y_i is a continuous health outcome. The RCI of y results from a relative concentration curve, which graphs on the x-axis the cumulative percentage of the sample, ranked by an indicator of socioeconomic position, such as living standards, or income beginning with the poorest. The y-axis then indicates the cumulative percentage of the health outcome corresponding to each cumulative percentage of the distribution of the socioeconomic indicator. Figure 1 provides an example of a relative concentration curve, where the health variable is malnutrition in Ghana in 2003. Figure 1 is constructed using data from the Ghana Demographic Health Survey in 2003 which was analysed in Van de Poel et al. (2007). It shows that the level of malnutrition accumulates faster amongst the poor than

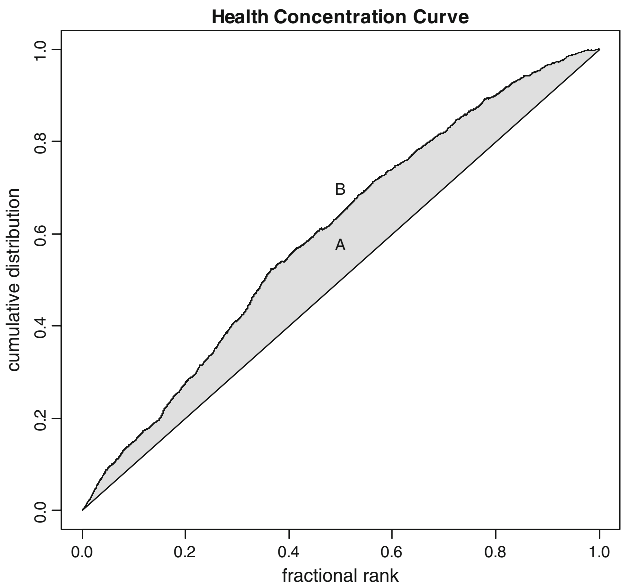

Fig. 1 The relative concentration curve: an example with malnutrition (Ghana 2003, *Source*: Demographic and Health Survey). The relative concentration index equals 2 × the area between the 45° line and the relative concentration curve = $A/(A + B)$

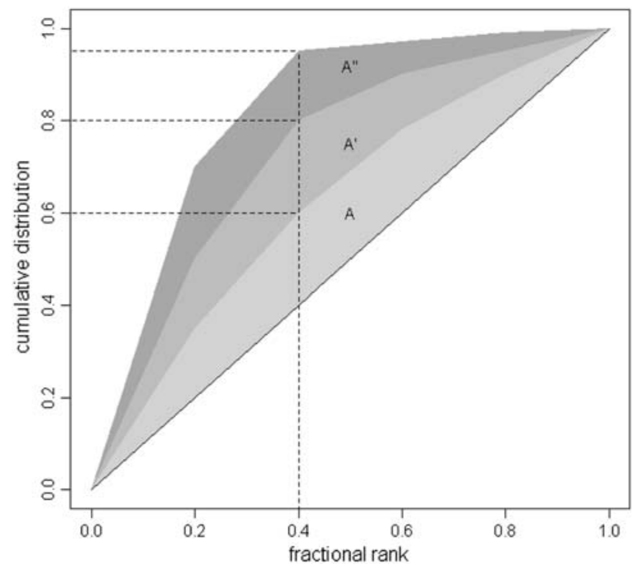

Fig. 2 Different relative concentration curves: an example with bad health. Section A demonstrates the lowest level of social inequality

amongst the better-off. For example, the poorest 40% (x-axis) of the population has nearly 60% of the malnutrition share. The RCI is defined as twice the area between the concentration curve, L(p) (the L stands for Lorenz curve), and the line of equality (the 45° diagonal from the bottom-left corner to the top-right). So, in the case where there is no income-related inequality, the RCI is zero.

The RCI takes a negative value when the curve lies above the line of equality, indicating disproportionate concentration of the health outcome among the poor, and a positive value when it lies below the line of equality (O'Donnell et al. 2008). If the health variable represents a 'bad' health state (e.g. malnutrition) a negative value of the RCI means ill health is higher among the poor.

Figure 2 shows a theoretical example of three different curves. The first one has an area A between the relative concentration curve and the diagonal (light grey), and the level of ill health (e.g. malnutrition) accumulates faster amongst the poor than amongst the better-off. The dotted line shows that the poorest 40% (x-axis) of the population has nearly 60% of the malnutrition share. The second curve has an area $A + A'$ between the concentration curve and the diagonal (and additional darker grey area of A') and the level of ill health (e.g. malnutrition) accumulates even faster amongst the poor than compared to the previous situation. The RCI will be higher as well (i.e. malnutrition is more disproportionately concentrated amongst the poor in this case). The dotted line shows that the poorest 40% (x-axis) of the population has now about 80% of the bad health share. Finally, a third curve has a concentration index which is even higher, getting closer to one (the additional dark grey area A'' results in an area which covers a good part of the whole upper triangle and twice this area is 1), and here 40% of the poorest have nearly 100% of the ill health share.

The value of the RCI normally varies between −1 and +1 but it is not bounded within the range of [−1, 1] if the health variable of interest takes negative, as well as positive values. Therefore, the health variable should be such that it is restricted to positive values. Further, the bounds of the RCI depend upon the mean (i.e. overall prevalence) of the indicator when applied to binary indicators (Wagstaff 2005). To avoid this problem, an alternative but related RCI that was recently introduced may be useful (Erreygers 2009). More formally, a RCI of y can be computed as twice the covariance of the health variable and a person's relative rank in terms of socioeconomic position, divided by the mean health according to Eq. 1:

$$\mathrm{RCI} = \frac{2}{\mu}\mathrm{cov}(y_i, R_i) \quad (1)$$

with y_i and R_i, respectively, the health status of the ith individual and the fractional rank of the ith individual in distribution of socioeconomic position (i.e. $R_i = 1/N$ for the poorest individual and $R_i = N/N$ for the richest); μ is the mean of the health of the sample and cov denotes the covariance. Notice that this assumes that the welfare variable is continuous. Methods for extending the calculation of RCI in cases where the welfare measure is categorical are available as well (Chen and Roy 2009). Furthermore, methods are available for calculating the RCI using weighted data (O'Donnell et al. 2008).

At the moment most of the software for inequality in health that is used in practice is written in Stata code (O'Donnell et al. 2008), e.g. the "concindexi" command has the Kakwani and Wagstaff covariance formulae, as outlined with standard errors. There is no implementation available at the moment that runs entirely on free software, allowing for the calculation of the RCI and associated confidence intervals. The methods described above are now implemented in an R package called decomp (available from the authors upon request). The R program is free of charge (http://www.r-project.org). This makes it an attractive option, especially in lower income countries. In an Appendix at the end of this paper, the use of the code is exemplified. After installing the package decomp into R, the code can be copied and pasted into R and immediately used after having adapted the variables to the users' needs.

Finally, it can be noted that the concentration index provides an indication of health inequalities but it is important that public health science moves from purely describing health inequalities towards quantifying the importance of different determinants of health inequality. As mentioned earlier, such a "decomposition" of health inequalities will be the topic of an upcoming paper in this journal.

Acknowledgments This paper resulted from cooperation initiated through meetings organized by the WHO European Office for Investment for Health and Development, Venice, Italy.

Appendix

We use a Demographic Health Survey dataset from Kenya (2003) for illustrating this new R-package. Commands and output are in courier font. Commands are indicated by a prompt (>).

Start R, and load the decomp package (which includes the 'kenya' dataset):

```
> library(decomp)
```

The command "RCI" calculates the concentration index for stunting and its confidence interval.

```
> kenya.stunting.CI <- RCI(kenya$wealth, kenya$stunting, wt = kenya$popweight)
> summary(kenya.stunting.CI)
Call:
RCI(x = kenya$wealth, y = kenya$stunting, wt = kenya$popweight)
Health Concentration Index:
 -0.1153306
Variance:
0.0001241775
95% Confidence Interval:

-0.1371715 -0.09348977
```

The next command will return a concentration plot.

```
> plot(kenya.stunting.CI)
```

Since the stunting variable contains both negative and positive values, the calculated RCI is not bounded between [-1,1]. This can be corrected in a number of ways, using the function 'correctSign'. One way is to impute a value of 0 for all negative values, as per (Van de Poel et al. 2007)

```
> kenya$stunting.corrected <- correctedValue(correctSign(kenya$stunting,
shift=FALSE))

> kenya.correctedstunting.CI <- RCI(kenya$wealth, kenya$stunting.corrected,
wt = kenya$popweight)
> summary(kenya.correctedstunting.CI)
Call:
RCI(x = kenya$wealth, y = kenya$stunting.corrected, wt = kenya$popweight)

Health Concentration Index:
 -0.08680988

Variance:
 5.265099e-05

95% Confidence Interval:
 -0.1010316 -0.07258818
```

References

Chen Z, Roy K (2009) Calculating concentration index with repetitive values of indicators of economic welfare. J Health Econ 28:169–175

Erreygers G (2009) Correcting the concentration index. J Health Econ 28:504–515

Harper S, Lynch J (2007) Commentary: using innovative inequality measures in epidemiology. Int J Epidemiol 36:926–928

Mackenbach JP, Stirbu I, Roskam AJR, Schaap MM, Menvielle G, Leinsalu M, Kunst AE (2008) Socioeconomic inequalities in health in 22 European countries. N Engl J Med 358:2468–2481

O'Donnell O, van Doorslaer E, Lindelöw M, World B (2008) Analyzing health equity using household survey data: a guide to techniques and their implementation. World Bank, Washington, DC

Smits J, Monden C (2009) Length of life inequality around the globe. Soc Sci Med 86:1114–1123

Van de Poel E, Hosseinpoor AR, Jehu-Appiah C, Vega J, Speybroeck N (2007) Malnutrition and the disproportional burden on the poor: the case of Ghana. Int J Equity Health 6:21

Wagstaff A (2005) The bounds of the concentration index when the variable of interest is binary, with an application to immunization inequality. Health Econ 14:429–432

Wagstaff A, Paci P, vanDoorslaer E (1991) On the measurement of inequalities in health. Soc Sci Med 33:545–557