

Outline methodology for OPSN composite indicators

We attempted to follow the approach laid out in the OECD's *Handbook on constructing composite indicators: methodology and user guide*¹.

Step 1: Theoretical framework

This is an exploratory project that set out to explore how far public data about mental health services could be turned into useful public information for people using those services. A key requirement was that data was publically available. The panel agreed four composite indicators:

1. How well is my GP looking after my physical health needs?
2. What is the likelihood of getting access to the right psychological therapies, and what is the outcome if I do?
3. Am I more or less likely than average to be prescribed anti-depressants?
4. How well am I supported to live well with my condition?

Clinical Commissioning Groups (CCGs) were selected as the unit of analysis.

Step 2: Data selection

A provisional long list of candidate indicators was developed for each composite indicator and was reviewed by the panel. Each indicator was checked for completeness and for showing reasonable variability. The panel indicated that they were keen for some additional indicators to be included in future versions of the composite when the data was available at CCG level.

Step 3: Imputation of missing data

As the purpose was to explore new ways of presenting data to the public, only indicators which were relatively complete were included as candidate indicators. CCGs with missing data were dropped from this analysis. These should be revived in future revisions.

Step 4: Multivariate analysis

Principal component analysis (PCA)² was used to explore how different indicators change in relation to each other and how they are associated.

Cronbach Coefficient Alpha³ was used to measure the internal consistency in the set of individual indicators, i.e. how well they describe a common construct.

Step 5: Normalisation

The individual metrics are on a range of scales. Therefore, normalisation is used to render the variables comparable. In step two and four we noted that most indicators contained a large number of outliers. Therefore ranking was chosen as a normalisation method as it is the least sensitive to outliers.

¹<http://www.oecd.org/std/42495745.pdf>

²<https://onlinecourses.science.psu.edu/stat505/node/49>

³<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4205511/>

Step 6: Weighting and aggregation

The panel was asked to weight each individual indicator and suggested that all should be equally weighted. Step four showed no highly correlated sub scale within a composite indicator and therefore there was minimal need to adjust for correlation within the indicators. The mean rank was used as the aggregation technique.

Step 5: Uncertainty and sensitivity analysis

Most of the individual indicators may have uncertainty (measured by confidence intervals) around them. For instance, the IAPT referrals: Rate (quarterly) per 100,000 population aged 18+ for a particular CCG might be 517 with 95 per cent confidence intervals ranging from 447 to 558. Therefore, rather than simply ranking the composite indicators, we adopted the approach outlined in Marshall and Spiegelhalter (1998)⁴ to allow for the uncertainty of the individual indicators. The R software has an add on library called **Compind** which is designed to facilitate the development of composite indicators in line with the OECD methodology. We used this package to produce one million random drawings for each indicator with its confidence intervals and from its underlying distribution. We then counted the number of times a CCG came either in the top 10 or the bottom 10 overall ranking of CCGs. The CCG that came in the top 10 (or bottom 10) more than five per cent of the time were considered either 'above expected' or 'below expected'.

The OECD methodology outlines a number of methods for normalisation and aggregation of individual indicators into a compound indicator. An alternative development team could legitimately argue for different approaches at each stage.

Then to understand sensitivity to our choices, we repeated the approach outlined in Marshall & Spiegelhalter (1998) with all the different normalisation and aggregation methods available in the **Compind** library. This indicated that the groupings into 'above expected' or 'below expected' were quite sensitive to the choice of ranking and aggregation method. No CCG was a convincing outlier across all methods.

Professor Simon Jones, New York University
David Mullett, University of Surrey
Alex Kafetz, ZPB Associates
Roger Taylor, The RSA Open Public Services Network

November 2015

⁴ Marshall, E.C., Spiegelhalter, D.J.1998. Reliability of league tables of in vitro fertilisation clinics : retrospective analysis of live birth rates. *BMJ*, 316, pp.1701–1705. <http://www.ncbi.nlm.nih.gov/pubmed/9614016>