

BMJ Open Do alternative weighting approaches for an Index of Multiple Deprivation change the association with mortality? A sensitivity analysis from Germany

Florian Schederecker,^{1,2} Christoph Kurz,¹ Jon Fairburn,³ Werner Maier¹

To cite: Schederecker F, Kurz C, Fairburn J, *et al*. Do alternative weighting approaches for an Index of Multiple Deprivation change the association with mortality? A sensitivity analysis from Germany. *BMJ Open* 2019;9:e028553. doi:10.1136/bmjopen-2018-028553

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2018-028553>).

Received 13 December 2018
Revised 11 April 2019
Accepted 13 June 2019



© Author(s) (or their employer(s)) 2019. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Institute of Health Economics and Health Care Management, Helmholtz Zentrum München (GmbH), Neuherberg, Germany

²IBE - Institute for Medical Information Processing, Biometry and Epidemiology, LMU Munich, Germany

³Business School, Staffordshire University, Stoke-on-Trent, UK

Correspondence to
Dr. Werner Maier;
werner.maier@helmholtz-muenchen.de

ABSTRACT

Objectives This study aimed to assess the impact of using different weighting procedures for the German Index of Multiple Deprivation (GIMD) investigating their link to mortality rates.

Design and setting In addition to the original (normative) weighting of the GIMD domains, four alternative weighting approaches were applied: equal weighting, linear regression, maximization algorithm and factor analysis. Correlation analyses to quantify the association between the differently weighted GIMD versions and mortality based on district-level official data from Germany in 2010 were applied (n=412 districts).

Outcome measures Total mortality (all age groups) and premature mortality (<65 years).

Results All correlations of the GIMD versions with both total and premature mortality were highly significant ($p<0.001$). The comparison of these associations using Williams's t-test for paired correlations showed significant differences, which proved to be small in respect to absolute values of Spearman's rho (total mortality: between 0.535 and 0.615; premature mortality: between 0.699 and 0.832).

Conclusions The association between area deprivation and mortality proved to be stable, regardless of different weighting of the GIMD domains. The theory-based weighting of the GIMD should be maintained, due to the stability of the GIMD scores and the relationship to mortality.

INTRODUCTION

Indices of deprivation are increasingly being used to investigate health and, in some countries, as tools of public policy.^{1–5} Therefore, it is important that these indices are transparent and rigorous in their construction so that confidence and understanding in their use are maintained.

In the 2000s, a series of deprivation indices with a multidimensional structure were introduced in the UK. These 'Indices of Multiple Deprivation' (IMDs) have been updated regularly ever since.⁶ The domains of deprivation were identified from the literature and were a result of the availability of data at the time.

Strengths and limitations of this study

- There is only limited literature on the application of different weighting approaches of deprivation indices—this study adds to that body of work.
- Our study provides an overview of established weighting approaches for deprivation indices used in Europe.
- Sensitivity testing of deprivation indices is particularly important as there seems to be no gold standard.
- We compare a broad range of normative and empirical weighting approaches for the domains of an Index of Multiple Deprivation.
- Limitations of the study concern the selection of weighting methods resulting from restricted data access at regional level.

A key aspect to consider when constructing such indices is the weighting and consolidation of the different deprivation domains that produce the final overall index.

Transparency and availability of data used in the indices mean that indicators and weightings can be adapted to particular demands by researchers. Adaptation may be needed, for example, to prevent autocorrelation effects where a component of the index is also related to the independent variable under consideration.

An IMD for Germany has been developed based on the methodology according to Noble *et al*.⁶ It was first applied in the German federal state of Bavaria ('Bavarian Index of Multiple Deprivation', BIMD) and subsequently as a nationwide IMD ('German Index of Multiple Deprivation', GIMD).^{7,8} For the construction of the German deprivation indices, domains from the British IMDs were partly used (eg, income and employment), and additional domains for social capital and municipal revenue were introduced. The GIMD includes both aspects, material deprivation (eg, income) as well as social deprivation (eg, social capital).

The GIMD has been used repeatedly for analyses regarding the relationship between area deprivation and morbidity, mortality and healthcare provision in Germany, and a persistent positive association has been shown between area deprivation and health outcomes.^{9–11}

One crucial point in building IMDs involves the weighting of the different deprivation domains. So far, weightings of IMDs have been conducted mainly by analysing literature on multiple deprivation and based on expert consultation.¹² Regarding the domain weights of the English IMD, alternative empirical weightings were carried out by C Dibben, which led to a recommendation of adjustment of the weights.¹³ However, this did not yield an alteration in the weighting of subsequent IMDs, as user surveys 'did not reveal significant support for moving to new weights',¹² and consisted only of two different empirical methods.

Besides the IMDs in the UK and Germany, several alternative approaches to the development and weighting of deprivation indices have been developed in other European countries¹⁴ as well as non-European countries.^{15 16} These approaches consist of a variety of (empirical) weighting approaches, which have not been applied to the British IMDs. However, it seems that almost all the approaches to weight deprivation indices are based on single methods, and sensitivity analysis regarding the application of different methods to a specific deprivation index has not been done. Additionally, literature regarding the application of different weighting procedures to a deprivation index is lacking.

As the GIMD was weighted by experts following the model of the British IMDs, we conducted a sensitivity analysis for the domain weighting of the GIMD following the example of Dibben *et al.*¹³ The aim of this study was to test the stability of the GIMD to different weighting approaches by conducting correlation analyses with mortality as a key health outcome. We decided to examine several alternative weighting approaches for the domains of the GIMD by stepwise comparison:

1. From a literature review, we obtained an overview of weighting approaches for deprivation indices in Europe and selected methods that can be used for alternative weighting approaches to the domains of the GIMD.
2. Regarding the weighting of the domains and the distribution of the GIMD scores, we analysed the results of the different weighting approaches and compared them with each other.
3. We compared the associations of these new versions of the GIMD with total mortality (all age groups) and premature mortality in Germany (<65 years) in order to conduct a sensitivity analysis concerning the different approaches.
4. Finally, we identified the weighting set that maximises the association between the GIMD and mortality.

A conceptual distinction between the different weighting methods was established with the identification of normative and empirically based approaches.

METHODS

Data for the statistical analysis

In order to construct different GIMD versions using alternative weighting approaches, we used regional data from the original GIMD from 2010 (GIMD 2010) for the domain and composite scores of the 412 districts in Germany.¹⁷ For the construction of the original GIMD, Maier *et al.*^{7 17} standardised nine deprivation indicators and assigned them to seven deprivation domains, which represent different dimensions of deprivation: income, employment, education, environment, security, municipal revenue and social capital (online supplementary 1). Each district is provided with a deprivation score for every single domain. The domain score is a statistical measure for the extent of area deprivation in a regional unit. The higher the deprivation within a district, the higher the domain score for the district. Subsequently, the domain scores are weighted based on a theoretical foundation and expert consultation and summed for an overall deprivation score for every district. For further details, see Maier *et al.*^{7 17}

Regarding an analysis of the relationship between area deprivation and both total mortality and premature mortality, we used raw mortality data and population data from 2010 at the district level, derived from the German Federal Statistical Office.¹⁸ The districts were identified by official district code numbers. Using the mortality and population data, we indirectly calculated standardised mortality rates (SMRs) for both total mortality (SMR 'total') as well as premature mortality (SMR 'premature'). This was necessary to compare districts because of their highly varying population size.¹⁹ For details on the calculation of the SMR, see online supplementary 2.

We used the variable 'available living space per inhabitant' from the German Federal Statistical Office from 2010¹⁸ as a proxy of deprivation. We reversed the polarity of its values and thus make it more comparable to the GIMD scores.

Methods for the weighting and methods for the statistical analysis

Additional to the original weighting of the GIMD 2010, we decided to use four methods for the weighting of the GIMD domains found in a literature review (table 1). We searched relevant literature in the databases PubMed and Embase (eg, keywords used in PubMed: (deprivation OR deprived) AND (index OR indices) AND (area* OR region* OR neighborhood OR neighbourhood), limits: English OR German OR French OR Italian OR Spanish).

Besides the equal weighting of the domains, we used two commonly used empirical methods and an additional greedy maximization algorithm method. The purpose of the empirical approaches was to extract relative weights for the domain scores from an empirical dataset. The extracted coefficients of the methods were used as relative weights for the domain scores, which should sum to 1 (or 100%), before the summation of the domains to an overall deprivation score.

Table 1 Overview of identified weighting methods through a literature review: characteristics and evaluation of the methods (abbreviations in brackets)

Weighting methods	Normative weighting of the domains/indicators		Empirical weighting of the domains/indicators				Confirmatory factor analysis (CFA)	Exploratory factor analysis (EFA)	Revealed preferences	
	Equal weighting of domains/indicators	Expert weighting of the domains/indicators	Theory-based weighting of the domains	Logistic regression analysis	Principal components analysis (PCA)	Bayesian factor analysis				
Description and weighting of the indicators / domains	► Equal weighting of the domains/indicators.	► Weighting of the domains/indicators according to expert opinion.	► Weighting of the domains: derived from research literature on multiple deprivation and social exclusion and by consultation process.	► Deprivations: proxy as dependent, coefficients as relative weights of the domains.	► Extraction of factors from indicators. Factors used as deprivation index/indices. Factor loadings as relative weights of the indicators. Assume the existence of an unmeasured unifying concept, but make no prior judgements as to what that is.	► Weights derived directly from the data. ³⁸	► Weights derived directly from the data. ³⁸	► Statistical model. Derivation of number of factors by model fit. ³⁹	► Number of factors derived from research literature. Factor loadings as relative weights of the indicators.	
Construction index	► Additive score of the indicators or domains.	► Additive score of the weighted indicators or domains.	► Different weighting of the indicators, according to the individual relevance.	► Weights derived from theory and research literature. ⁶	► Derives outcome specific weights from the data. ³⁸	► Easy handling, often used and robust approach. ³⁸⁻⁴¹	► Suited for analysis of small area units. ⁴²	► Exploration of latent dimensions without foreknowledge. ²⁸	► Dimensions of deprivation derived from theory and set a priori. ⁴¹	
Selected advantages of the methods	► Simplest solution to aggregation of indicators to an overall score. ³⁸	► Based on expert knowledge. ^{38,39}	► Selected weights dependent on the quality of the research from the literature.	► Removal or addition of variables can alter the coefficients significantly. ³⁸	► Derived coefficient dependent on the data quality.	► All variables load on all factors. ⁴¹	► At least weekly informative prior information is required.	► Restricted temporal comparability. ³⁸	► Elaborate procedure: theoretical knowledge and conceptualisation required. ²⁹	
Selected disadvantages of the methods	► Unintentional, implicit weighting of the domains of multiple deprivation possible (owing to the number of indicators included per dimension). ⁶	► Arbitrary weighting possible, because of subjective decisions. ³⁸	► Large or multiple variables bias possible.	► Removal or addition of variables can alter the coefficients significantly. ³⁸	► Transferability to the population limited because of explanation of the sample variance. ²⁸	► Reduction to one factor does not consider the multidimensionality of deprivation. ⁴¹	► One factor does not consider the multidimensionality of deprivation. ⁴¹	► Different interpretability of the results, several decisions required. ²⁹	► Several decisions required regarding covariance structure and the method of parameter estimation. ²⁹	
Selected examples	43	23	38,12	13	44	42	45	46	13	

- Original weighting of the domains of the GIMD through theoretical foundation and expert opinion according to Maier et al.⁸ For weights used, see online supplementary 1.
- Equal weighting of the seven GIMD domains; thus, each domain weighted with 1/7. This approach was originally used for deprivation indices by Carstairs and Townsend.^{20 21} To date, this approach is still used for deprivation indices consisting of just single deprivation indicators.^{21 22} For this approach, an equal importance of all deprivation indicators is assumed. In our analysis, we transferred this approach to the domain level.
- Weighting of the domains by the coefficients of a linear regression analysis with a proxy for deprivation ('available living space per inhabitant') as the dependent variable and the GIMD domains as the independent variables. We had to choose a dependent variable for the linear regression that had not been used for the construction of the GIMD domains and could be considered as an indicator of deprivation.¹³ Townsend, Carstairs and Jarman considered overcrowding of living space as an indicator of deprivation.^{20 21 23} We assumed that the availability of living space per inhabitant in an area could act as a proxy for area deprivation: the more deprived the area, the less living space is available per inhabitant.²⁴⁻²⁶ For this approach, we calculated the absolute value of the regression coefficients and then used them as relative weights for the specific domains. Subsequently, the weighted domain scores were summed to an overall score. Linear models for the extraction of weights for a deprivation score have already been conducted in several studies.^{13 27 28} Because of the normal distribution of the dependent variable, we conducted an ordinary least square regression.
- Weighting of the GIMD domains using a greedy maximization algorithm (Kurz C, Maier W, Rink C. A greedy stacking algorithm for model ensembling and domain weighting. Working paper. 2019). This yields weights for the domains close to the maximum possible correlation between the GIMD 2010 and mortality as a relevant outcome of deprivation (online supplementary 3). The weighted domain scores of the GIMD were then added together to an overall index for both total mortality and premature mortality. This addition to the methods of the literature search aimed to extract weights for the maximum Spearman correlation between GIMD and mortality and can thus be seen as an outcome-specific approach with the independent variable mortality. Complete circularity was present because mortality had already been used for the extraction of the weights. In contrast, the other methods could be seen as general weighting approaches for deprivation indices.
- Weighting of the domains according to the results of an exploratory factor analysis (EFA). We chose a principal axis factoring (PAF) approach for the extraction of the factors. PAF is a commonly used extraction method for factor analysis and requires no specific distribution of the entered variables. This non-parametric approach was necessary because of the exponentially

transformed domains.²⁹ A priori, we specified the extraction of one factor (as a latent factor, measuring 'multiple deprivation') out of the seven domains. The absolute values of the factor loadings of the different domains were used as relative weights for the domains. Again, the weighted domains were added together to an overall deprivation score.

Correlation analysis and statistical software

Subsequently, we performed a sensitivity analysis of the newly weighted GIMD versions. We conducted correlation analyses in order to calculate the relationship between the different GIMD versions and both total as well as premature mortality (in terms of SMRs) and compared their results. For the analysis, we used Spearman's rank correlation coefficient (ρ) as a robust approach. This was required, in our opinion, as the GIMD score could be interpreted as an ordinal variable because of the ranking of the districts during the generation of the domain scores.^{7 30} Correlation analyses were each performed with a GIMD version and both total mortality and premature mortality. We also tested for significance of these bivariate correlation coefficients at an α -level of 5%.³¹ For comparison of the bivariate correlations among each other, we performed t-tests for paired correlations. For this, we used Williams's t-test for the comparison of correlations out of dependent samples.³² We compared two correlation coefficients in terms of both total and premature mortality at an α -level of 5%. For the statistical analysis, we used the Software R 3.2.3.³³

Patient and public involvement

Patients and/or public were not involved in this study.

RESULTS

Population size of the districts and estimation of the SMRs

The size of the population of the 412 districts varied with median size of 139 188 inhabitants, interquartile range (IQR) of 130 170 persons, minimum size of 33 944 and maximum size of 3 460 725 persons. Raw mortality of the 412 districts varied with median of 1522 death cases, IQR of 1347 cases, minimum of 413 cases and maximum of 32 234 cases. Qualifying date of the data was 31 December 2010. We estimated total mortality by calculating 'SMR_{total}' for the districts with a mean of 1.0175 (SE: 0.004) and premature mortality 'SMR_{premature}' with a mean of 1.0165 (SE: 0.004).

Weights of the domains of the alternative approaches

An overview of the identified weighting methods for deprivation indices is given in table 1. Alongside a description of the weighting and the construction of the deprivation indices, we offered selected advantages and disadvantages of the methods. This was completed with selected examples. From this table, we chose four approaches additional to the original weighting of the GIMD.

Table 2 Weighting of the domains of the GIMD through different weighting approaches, values in percentage points

Deprivation domains/methods of domain weighting of the GIMD	Income	Employment	Education	Municipal revenue	Social capital	Environment	Security
Original weighting	25.00	25.00	15.00	15.00	10.00	5.00	5.00
Equal weighting	14.29	14.29	14.29	14.29	14.29	14.29	14.29
Linear regression	4.47	21.68	15.42	30.25	11.45	14.65	2.09
Maximization algorithm (total mortality)	18.23	20.67	1.04	21.90	28.26	4.62	5.28
Maximization algorithm (premature mortality)	18.85	48.93	0.31	15.98	10.73	0.50	4.70
Factor analysis	23.09	18.99	8.97	21.72	20.08	5.86	1.28

Original weighting: weighting according to Maier *et al.*⁸

Equal weighting: every domain gets equal weighting (1/7=0.1429).

Linear regression: weighting of the domains with regression coefficients with a deprivation proxy as dependent and domains as independent variables.

Maximization algorithm: weighting of the domains in order to achieve the maximum Spearman correlation between overall index and mortality.

Factor analysis: weighting of the domains with loadings from principal axis factoring.

GIMD, German Index of Multiple Deprivation.

We found considerable differences between the domain weights resulting from the different approaches (table 2).

The weights for employment deprivation showed the largest variation with a range of 34 percentage points. The other deprivation domains showed a range of at least 13 percentage points. Educational deprivation within the maximization algorithm and income deprivation within the linear regression showed very small weights compared with the weights of the original GIMD 2010. Municipal revenue deprivation resulted in a weight twice as high as the original weight within the linear regression. Concerning the algorithm, the weight for social capital deprivation was three times the original weight. Concerning premature mortality, the weight for employment deprivation was twice as high as the original weight for the GIMD. Deprivation domains for social capital

and district income showed constantly higher weights for the empirical approaches compared with the two normative methods. The different GIMD versions revealed different distributions of the overall deprivation scores (table 3).

Assumptions for the linear regression were generally met, and the model had significant explanatory power (adj. $R^2=0.33$). Five of the seven domains showed a significant effect on the deprivation proxy. Heterogeneity was present; thus, we presented robust SEs. Additionally, we provided tests of the assumptions of the linear regression model (online supplementary 4). The factor analysis generally had significant explanatory power ($\chi^2: 584.65$, $p<0.0001$), but showed low reliability (Tucker-Lewis index=0.50) and an RMSEA of 0.32 with tight CIs (0.30 to 0.34) indicated that this one factor was not a good fit to the data (online supplementary 5).

Table 3 Descriptive results of the weighted indices, information on GIMD scores

	Original weighting	Equal weighting	Linear regression	Maximization algorithm SMR 'total' (SMR 'premature')	Factor analysis
No of districts	412	412	412	412 (412)	412
Mean	21.81	21.81	21.81	21.81 (21.81)	21.81
Median	18.80	19.97	19.34	17.05 (16.49)	18.17
SD	12.73	10.34	10.98	15.61 (17.09)	14.24
Minimum	2.04	2.29	2.11	1.48 (0.92)	1.33
Maximum	70.98	55.69	67.67	85.91 (91.14)	79.86

Original weighting: weighting according to Maier *et al.*⁸

Equal weighting: every domain gets equal weighting (1/7=0.1429).

Linear regression: weighting of the domains with regression coefficients with a deprivation proxy as dependent and domains as independent variables.

Maximization algorithm: weighting of the domains in order to achieve the maximum Spearman correlation between overall index and both total mortality (SMR 'total') and premature mortality (SMR 'premature' in brackets).

Factor analysis: weighting of the domains with loadings from principal axis factoring.

GIMD, German Index of Multiple Deprivation; SMR, standardised mortality ratio.

Table 4 Spearman's rank correlation coefficients for the association between the versions of the GIMD and both premature and total mortality

Methods of the domain weighting of the GIMD	Total mortality (SMR 'total')	Premature mortality (SMR 'premature')
Original weighting	0.578* [0.506 to 0.642]	0.767* [0.718 to 0.808]
Equal weighting	0.535* [0.459 to 0.604]	0.699* [0.641 to 0.750]
Linear regression	0.564* [0.492 to 0.629]	0.738* [0.685 to 0.784]
Maximization algorithm	0.615* [0.547 to 0.676]	0.832* [0.794 to 0.864]
Factor analysis	0.598* [0.529 to 0.661]	0.772* [0.724 to 0.813]

Original weighting: weighting according to Maier *et al.*⁸

Equal weighting: every domain gets equal weighting (1/7=0.1429).

Linear regression: weighting of the domains with regression coefficients with a deprivation proxy as dependent and domains as independent variables.

Maximization algorithm: weighting of the domains in order to achieve the maximum Spearman correlation between overall index and mortality.

Factor analysis: weighting of the domains with loadings from principal axis factoring.

*p<0.001, bootstrapped (10000-fold) 95% CIs in square brackets.

GIMD, German Index of Multiple Deprivation; SMR standardised mortality ratio.

DISCUSSION

The central objective of the study was to explore whether alternative weighting approaches had an influence on the relationship between area deprivation and mortality when applied to the GIMD. Thereby, different weighting methods were selected if they were, on the one hand, applicable to the domain-based construction of the GIMD and, on the other hand, seemed feasible in the course of an application of a multimethodical approach. The four different methods were applied to the weighting of the domains of the GIMD 2010. The selected approaches and the original method were compared concerning both the weighting of the domains of the GIMD and the relationship between GIMD and mortality.

There was little evidence in the literature concerning the application of different weighting methods for multidimensional deprivation indices. However, a summary of different weighting approaches and their classification was presented by Noble *et al.*⁶ They briefly assessed the specific procedures of the methods (eg, empirical approaches) and were in favour of a weighting driven by literature considerations on multiple deprivation. Regarding the application of empirical weighting approaches for the English IMD 2004, we want to emphasise Dibben's work.¹³ He recommended new weights for the domains of the IMD, as the empirical weighting approaches indicated a higher weighting of the health domain and a lower weighting of the employment domain. However, this suggested swapping of weights was not eventually applied to the subsequent versions of the English IMD. The maintenance of the weights was justified by a consultation of IMD users and stable results of the IMD with either existing or suggested weights.¹²

In this study, we pursued a multimethodical approach for the weighting of the GIMD, including empirical methods. Owing to the different inherent intentions of the selected methods, we integrated the approaches as follows:

1. *Normative approaches*: the original weighting of the domains according to Maier *et al.*^{7,17} through theory and experts' opinion. We used the term 'normative' because weights for the domains must be selected a priori subjectively before they can be validated with data.
2. *Specific empirical approaches*: concerning the maximization algorithm with the dependent variable mortality, a weighting of the domains has been sought that was in line with the relationship between area deprivation and both total mortality and premature mortality and should maximise the correlation between them.
3. *General empirical approaches*: in contrast to the specific empirical approaches, the weighting of the domains was realised according to the results of a linear regression model or according to a factor analysis to generate generally applicable indices, which can also be used for the analysis of other health outcomes.

A further distinction of the methods can be made regarding their conceptual aspects. Factor analysis and principal components analysis are unsupervised methods

Results of the statistical analysis

Correlation analysis between the differently weighted deprivation indices and mortality showed different results (table 4).

Deprivation indices, domains weighted by the maximization algorithm, showed the maximum correlation with total mortality ($p=0.615$) and premature mortality ($p=0.832$). Correlations between the original GIMD and both total and premature mortality were $p=0.578$ and 0.767, respectively. Correlations between the equally weighted GIMD and mortality were the lowest with $p=0.535$ and 0.699. All correlations were significant concerning both total and premature mortality ($p<0.001$). Additionally, bivariate correlations between all indices were significant (p between 0.86 and 0.98).

Pairwise comparisons of the correlation coefficients with Williams's t-tests showed a differentiated result (see online supplementary 6). Almost every pairwise difference in the correlation coefficients was significant at the 5% α -level. One exception was the difference in the coefficients between the original GIMD and the GIMD weighted by linear regression concerning total mortality. The other deviation was the difference between the original GIMD and the GIMD weighted by factor analysis concerning premature mortality. The difference was not significant, neither one-sided nor two-sided. Maximum correlation coefficients of the GIMD, weighted by the algorithm, differed significantly from all the correlation coefficients of the other methods regarding both total and premature mortality. When we corrected for the multiple comparison of the difference of the correlation between the GIMD versions, there was a slight difference present in the significances (online supplementary 7).

that require no prior judgements and construct deprivation solely based on the domain knowledge. On the other hand, linear regression and the maximization algorithm are supervised or predictive methods considering deprivation based on a specific proxy and assuming a relationship between this proxy and deprivation.

Assessment of the alternative weighting approaches

The high weighting of the deprivation domains income and employment of 50% altogether within the original GIMD was confirmed by the empirical weighting of the factor analysis approximately, as well as the weighting of the environment deprivation domain. Educational deprivation was weighted considerably lower by the factor analysis and algorithm than by the original GIMD. Deprivation domains for district income and social capital were constantly weighted much higher by the empirical approaches than by the approach of the original GIMD. The shift in the weighting of the domains can be explained by the data dependency of the empirical approaches and should be reviewed using alternative data. Should the higher weighting of the district income and social capital domains be confirmed, an adjustment in the domain weights could be considered. Perhaps those context variables have a higher relevance concerning area deprivation than expected by Maier *et al.*⁸

The low weighting of the deprivation domains of income by the linear regression and education by the maximization algorithm can barely be reconciled with existing evidence regarding the positive relationship of these two deprivation domains and mortality.^{34 35} The high weighting of the employment deprivation domain (49%) by the algorithm, concerning mortality, could reflect the high relevance of unemployment relating to premature mortality.

Relationship of the GIMD versions and mortality

Throughout the analysis, we could not find a weighting method that could be seen as superior compared with the other approaches or could even be recommended as a gold standard. Even though almost all GIMD weighting approaches differed significantly in their correlation with mortality, using only significance as a method of evaluation for the approaches seemed inappropriate. The correlation coefficient between the different GIMD versions was already very high ($p>0.89$), so that even small non-relevant differences could have produced significant results. All correlations of the GIMD versions with mortality were highly significant and showed rather small differences in respect to absolute values (p between 0.54 and 0.62). Since we conducted multiple paired t-tests, type-1 error inflation was present. In an additional analysis, we corrected for multiple testing with Benjamini and Yekutieli adjustment.³⁶ When we corrected for the correlation of the GIMD versions with mortality, the significance of the results did not change (table 4). When we corrected for the multiple comparison of the difference of the correlation between the GIMD versions (online

supplementary 6), there was a slight difference present in the significance (online supplementary 7).

The empirical weighting of the GIMD by an EFA represented an adequate alternative to the theory-based weighting of the domains, on account of the simple operability and the highly significant association of this GIMD version with mortality. Thereby, a general applicability of the GIMD for the analysis of implications for other health outcomes can be ensured, and the results of different datasets can be compared by model fit measures.²⁹ Despite the significant correlation, the application of equal weighting of the domains could be considered as obsolete, as this would produce an implicit weighting of the domains depending on the availability of indicators for each domain.⁶

Strengths and limitations of the study

Using a multimethodological strategy, we were able to cover a broad bandwidth of weighting approaches. As there seems to be no gold standard for weighting of deprivation indices, we recommend that sensitivity testing of the GIMD is particularly important. An equal weighting as well as an EFA for the weighting of multiple deprivation domains were carried out in this study for the first time. A factor analysis of the IMD domains was advised by Deas *et al.*²² but has not been implemented to date. Furthermore, we provided an outcome-specific weighting approach in the form of a greedy maximization algorithm: this method produced a domain weighting of the GIMD that maximised a specific measure concerning one health outcome (in this case, the correlation between GIMD and mortality). A transfer of the algorithm to other areas of interest is possible without difficulty but should be used mainly for orientation, which is possible concerning a selected measure, given a dataset.

Limitations of the study concerned the selection of weighting methods such as the revealed preferences or Bayesian factor analysis (cf table 1), which resulted from restricted data access at a regional level. Empirical methods are always data dependent and are restricted concerning a possible comparison over time, especially with the use of cross-sectional data. This could be addressed by using longitudinal data and would enable us to measure 'between variation' (ie, over different locations) to 'within variation' (ie, the same location over time).

Using correlation coefficients to evaluate the association between different GIMD versions and mortality does not necessarily imply a causal association between area deprivation and mortality. Additionally, overfitting is present by using the greedy maximization algorithm as a weighting approach, since it already yields the weights for the maximum correlation between the GIMD and mortality. However, there is reliability of using the GIMD to evaluate total and premature mortality, since the correlation between the GIMD and mortality is very stable over time (GIMD scores from 2006 to 2010 yield very similar correlations with mortality). Another point

was the lack of literature regarding the application of different weighting procedures. This limitation could partly be counterbalanced with the input of expert interviews. With regard to the linear regression, the selection of the deprivation proxy should be reconsidered *ex post*, as the use of the deprivation measure regarding living space per inhabitant showed a rather weak (yet significant) positive correlation with overall deprivation ($p=0.35$). This could be explained by the idea that, in less deprived cities such as Hamburg and Munich, there can be—in general—less available living space because of a very competitive housing market. So, there could be a partial negative correlation between deprivation regarding available living space and overall deprivation in some areas. Unfortunately, multidimensional proxies at district level were not available for Germany. We tested other measures like the overall gross domestic product (GDP) per district and the GDP per employed persons per district. They had a similar or lower correlation with the original GIMD as the living space variable, but using them had some major drawbacks. We understand, that the use of a one-dimensional proxy is a limitation in our work. However, given the very restricted variety of appropriate variables at the district level in Germany, the selection of this proxy was a pragmatic approach to test a weighting approach based on a linear regression.

We are aware that the stability of the GIMD could have also been tested by applying systematic changes to the weighting of the GIMD domains without using a framework of different weighting approaches. The correlation between some deprivation domains (eg, income or employment) is relatively high and thus any weighting scheme would likely give highly correlated results with mortality. A recent study from the UK showed that 94% of the variance in the English IMD could be explained by the income and employment domains alone, even though they had weights of 22.5% each in the overall index. The authors stated that even if the weights for the other domains had been zero, there would have been very little impact on the overall index.³⁷ Nevertheless, the aim of our study was to provide a conceptual framework of weighting approaches (normative and empirical) for an index of multiple deprivation and to combine the results of the literature search with a sensitivity analysis based on the GIMD.

CONCLUSION

The variation in the domain weights of the GIMD did not have a large measurable impact on the relationship between area deprivation and mortality. The correlation between the GIMD and both total mortality and premature mortality proved to be very stable, regardless of the application of the different weighting approaches and the resulting different sets of domain weights. The GIMD versions produced relatively stable results with regard to the central distribution measures of the overall scores (table 3).

The theory-based weighting of Maier *et al* can be interpreted *ex post* as more conservative than the empirical weighting approaches, as the weighting of the income and employment domains is relatively strong at 50% in contrast to the empirical methods. Nevertheless, a theory-based selection of domains seems to be more meaningful than an empirically based selection because the results of the empirical methods are restricted, as discussed above. The stability with respect to the scores and the relationship to mortality supports this advice. A modelling of the GIMD with a confirmatory factor analysis could be considered as a promising empirical approach with the prospect of temporal comparability in future studies.

Contributors FS and WM designed the study. FS performed the analysis and wrote the initial version of the manuscript. CK gave statistical advice, provided the greedy maximization algorithm and gave support preparing the manuscript. WM provided the German Index of Multiple Deprivation (GIMD), supervised the study and gave support preparing the manuscript. JF advised on deprivation indices and gave support preparing the manuscript. All authors read and approved the final manuscript.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Extra data regarding the GIMD is available by emailing Werner Maier (werner.maier@helmholtz-muenchen.de).

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

REFERENCES

1. Mackenbach JP, Kunst AE, Cavelaars AE, *et al*. Socioeconomic inequalities in morbidity and mortality in western Europe. The EU Working Group on Socioeconomic Inequalities in Health. *Lancet* 1997;349:1655–9.
2. Voigtlander S, Berger U, Razum O. The impact of regional and neighbourhood deprivation on physical health in Germany: a multilevel study. *BMC Public Health* 2010;10:403.
3. Siegel M, Mielck A, Maier W, *et al*. and Health: Do Income-Related Health Inequalities Vary by Small Area Deprivation? *Health economics* 2015;24:1523–30.
4. Zhang X, Cook PA, Lisboa PJ, *et al*. The effects of deprivation and relative deprivation on self-reported morbidity in England: an area-level ecological study. *Int J Health Geogr* 2013;12:5.
5. Allender S, Scarborough P, Keegan T, *et al*. Relative deprivation between neighbouring wards is predictive of coronary heart disease mortality after adjustment for absolute deprivation of wards. *J Epidemiol Community Health* 2012;66:803–8.
6. Noble M, Wright G, Smith G, *et al*. Measuring Multiple Deprivation at the Small-Area Level. *Environment and Planning A: Economy and Space* 2006;38:169–85.
7. Maier W, Fairburn J, Mielck A. [Regional deprivation and mortality in Bavaria. Development of a community-based index of multiple deprivation]. *Gesundheitswesen* 2012;74:416–25.
8. Maier W, Holle R, Hunger M, *et al*. The impact of regional deprivation and individual socio-economic status on the prevalence of Type 2 diabetes in Germany. A pooled analysis of five population-based studies. *Diabet Med* 2013;30:e78–e86.
9. Maier W, Scheidt-Nave C, Holle R, *et al*. Area level deprivation is an independent determinant of prevalent type 2 diabetes and obesity at the national level in Germany. Results from the National Telephone Health Interview Surveys 'German Health Update' GEDA 2009 and 2010. *PLoS One* 2014;9:e89661.

10. Schäfer T, Pritzkuleit R, Jeszenszky C, et al. Trends and geographical variation of primary hip and knee joint replacement in Germany. *Osteoarthritis Cartilage* 2013;21:279–88.
11. Jansen L, Eberle A, Emrich K, et al. Socioeconomic deprivation and cancer survival in Germany: an ecological analysis in 200 districts in Germany. *Int J Cancer* 2014;134:2951–60.
12. Smith T, Noble M, Noble S, et al. *The English Indices of Deprivation*. London: Department for Communities and Local Government, 2015.
13. Dibben C, Atherton I, Cox M, et al. *Investigating the impact of changing the weights that underpin the Index of Multiple Deprivation 2004*. London: Department for Communities and Local Government, 2007.
14. Fairburn J, Maier W, Braubach M. Incorporating Environmental Justice into Second Generation Indices of Multiple Deprivation: Lessons from the UK and Progress Internationally. *Int J Environ Res Public Health* 2016;13:750.
15. Cabrera-Barona P, Murphy T, Kienberger S, et al. A multi-criteria spatial deprivation index to support health inequality analyses. *Int J Health Geogr* 2015;14:11.
16. Pampalon R, Hamel D, Gamache P, et al. A deprivation index for health planning in Canada. *Chronic Dis Can* 2009;29:178–91.
17. Bauer H, Maier W. GIMD 2010 – Ein Update des, German Index of Multiple Deprivation*. *Berichte des Helmholtz Zentrums München* 2018 https://push-zb.helmholtz-muenchen.de/frontdoor.php?source_opus=54124&la=de (cited 4 Dec 2018).
18. Statistische Ämter des Bundes und der Länder [Internet]. The Regional Database Germany. 2016 <https://www.regionalstatistik.de/genesis/online/> (cited 25 Jun 2017).
19. Gregory IN. Comparisons between geographies of mortality and deprivation from the 1900s and 2001: spatial analysis of census and mortality statistics. *BMJ* 2009;339:b3454.
20. Carstairs V, Morris R. Deprivation and health in Scotland. *Health Bull* 1990;48:162–75.
21. Townsend P, Phillipson P, Beattie A. Health and Deprivation: inequality and the North (London, Croom Helm). *Links* 1988.
22. Deas I, Robson B, Wong C, et al. Measuring Neighbourhood Deprivation: A Critique of the Index of Multiple Deprivation. *Environment and Planning C: Government and Policy* 2003;21:883–903.
23. Jarman B. Underprivileged areas: validation and distribution of scores. *Br Med J* 1984;289:1587–92.
24. Atkinson J, Salmon C, Crampton P. NZDep2013 Index of Deprivation. 2014 <https://www.otago.ac.nz/wellington/otago069936.pdf> (cited 11 Jun 2018).
25. Boarini R, d'Ercole MM. Measures of Material Deprivation in OECD Countries. 2006 <http://www.oecd.org/social/soc/37223552.pdf> (cited 11 Jun 2018).
26. Eurostat. Housing conditions. 2017 http://ec.europa.eu/eurostat/statistics-explained/index.php/Housing_conditions#Housing_quality. C2.A0.E2.80.94.C2.A0overcrowding (cited 11 Jun 2018).
27. Ivaldi E, Testi A. Genoa Index of Deprivation (GDI): An index of material deprivation for geographical areas. *Social Indicators: Statistics, Trends and Policy Development* 2010;75.
28. Pernet C, Delpierre C, Dejardin O, et al. Construction of an adaptable European transnational ecological deprivation index: the French version. *J Epidemiol Community Health* 2012;66:982–9.
29. Thompson B. *Exploratory and confirmatory factor analysis: Understanding concepts and applications*: American Psychological Association, 2004.
30. Brown TA. *Confirmatory Factor Analysis for Applied Research*. New York: The Guilford Press, 2006.
31. Hollander M, Wolfe DA. *Nonparametric statistical methods*, 1973.
32. Williams EJ. *Regression analysis*. New York: wiley, 1959.
33. Team RC. *R: a language and environment for statistical computing*. Vienna, Austria, 2015.
34. Cutler D, Deaton A, Lleras-Muney A. The Determinants of Mortality. *Journal of Economic Perspectives* 2006;20:97–120.
35. Wolfson MC, Kaplan G, Lynch J, et al. Relation between income inequality and mortality: empirical demonstration. *West J Med* 2000;172:22–4.
36. Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics* 2001;29:1165–88.
37. Abel GA, Barclay ME, Payne RA. Adjusted indices of multiple deprivation to enable comparisons within and between constituent countries of the UK including an illustration using mortality rates. *BMJ Open* 2016;6:e012750.
38. Folwell K. Single measures of deprivation. *J Epidemiol Community Health* 1995;49(Suppl 2):S51–S56.
39. Gordon D. Census based deprivation indices: their weighting and validation. *J Epidemiol Community Health* 1995;49 Suppl 2:S39–S44.
40. Haase T, The PJ. Pobal HP Deprivation Index for Small Areas (SA): Introduction and Reference Tables. *Pobal* 2012.
41. Haase T, Pratschke J. *Deprivation and its Spatial Articulation in the Republic of Ireland*. Dublin: Area Development Management, 2005.
42. Mari-Dell'Olmo M, Martínez-Beneito MA, Borrell C, et al. Bayesian factor analysis to calculate a deprivation index and its uncertainty. *Epidemiology* 2011;22:356–64.
43. Slachová H, Tomášková H, Splíchalová A, et al. Czech socio-economic deprivation index and its correlation with mortality data. *Int J Public Health* 2009;54:267–73.
44. Meijer M, Engholm G, Grittner U, et al. A socioeconomic deprivation index for small areas in Denmark. *Scand J Public Health* 2013;41:560–9.
45. Testi A, Ivaldi E. Material versus social deprivation and health: a case study of an urban area. *Eur J Health Econ* 2009;10:323–8.
46. Haase T, Pratschke J, Gleeson J. All-Island Deprivation Index: Towards the development of consistent deprivation measures for the island of Ireland. *BORDERLANDS – The Journal of Spatial Planning in Ireland* 2012;2.