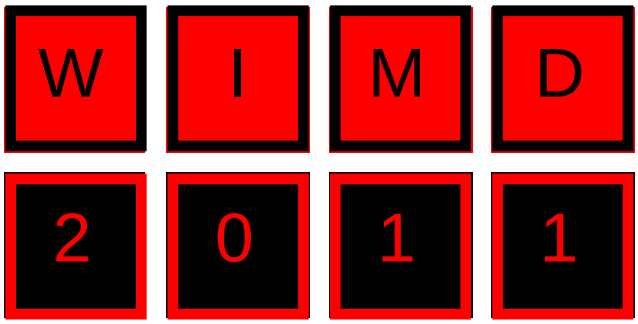
ystadegau ar gyfer cymru statistics for wales

Llywodraeth Cymru
Welsh Government



Welsh Index of Multiple Deprivation

Welsh Index of Multiple Deprivation 2011

Technical Report

Contents

1 Introduction

The Welsh Index of Multiple Deprivation (WIMD) 2011 is the official measure of **relative deprivation** for **small areas** in Wales and WIMD 2011: Child Index is the official measure of relative deprivation for children in Wales. Both Indexes are produced by the Welsh Government. The Indexes were developed as a tool to identify and understand deprivation in Wales, so that funding, policy, and programmes can be effectively focussed on the most disadvantaged communities. WIMD 2011 updates the WIMD 2008, and WIMD 2011: Child Index updates the WIMD 2008: Child Index, without imposing methodological changes, except where data provision makes this unavoidable.

This technical report describes how WIMD 2011 and WIMD 2011: Child Index was constructed and contains a full list of indicators and information about the indicators.

Deprivation is a wider concept than poverty. Poverty is usually means a lack of money, whereas deprivation includes a lack of the opportunities and resources to which we might expect to have access in our society, for example, good health or protection from crime. 'Multiple' deprivation therefore refers to the different types of deprivation that might occur. Eight domains (types) of deprivation are included in WIMD 2011: employment, income, education, health, community safety, geographical access to services, housing and physical environment. The same domains are used in WIMD 2011: Child Index, except employment is dropped. Each domain is made up of a number of indicators; in total there are 35 indicators used in WIMD 2011 and 26 indicators used in WIMD 2011: Child Index.

It should be noted that, due to the nature of the area-based data from which the Index is constructed, we do not know whether individuals are multiply deprived, or if different individuals are suffering different types of deprivation. At present, we are not able to link the data included in the Index in order to understand this. It is important to remember that a lack of deprivation is not the same as affluence. The least-deprived area is not necessarily the most affluent area in Wales.

The Index is produced as a set of ranks, with a rank of 1 assigned to the most deprived area. Ranks are a **relative** system of measurement; we can know which areas are more (or less) deprived than others, but not by how much. This is because of the way that the Index must be constructed. The construction of the Index is described in more detail below.

The ranks of the Index are calculated for each of the 1896 lower layer super output areas (**LSOA**s) of Wales. Although the *geographical* size of these **small areas** varies quite widely, the *populations* are intended to be roughly the same in each LSOA, with an average population of 1500 people. LSOAs were designed by the Office for National Statistics to have consistent population sizes and stable geographies, so that statistical comparisons of small areas over time can be carried out.

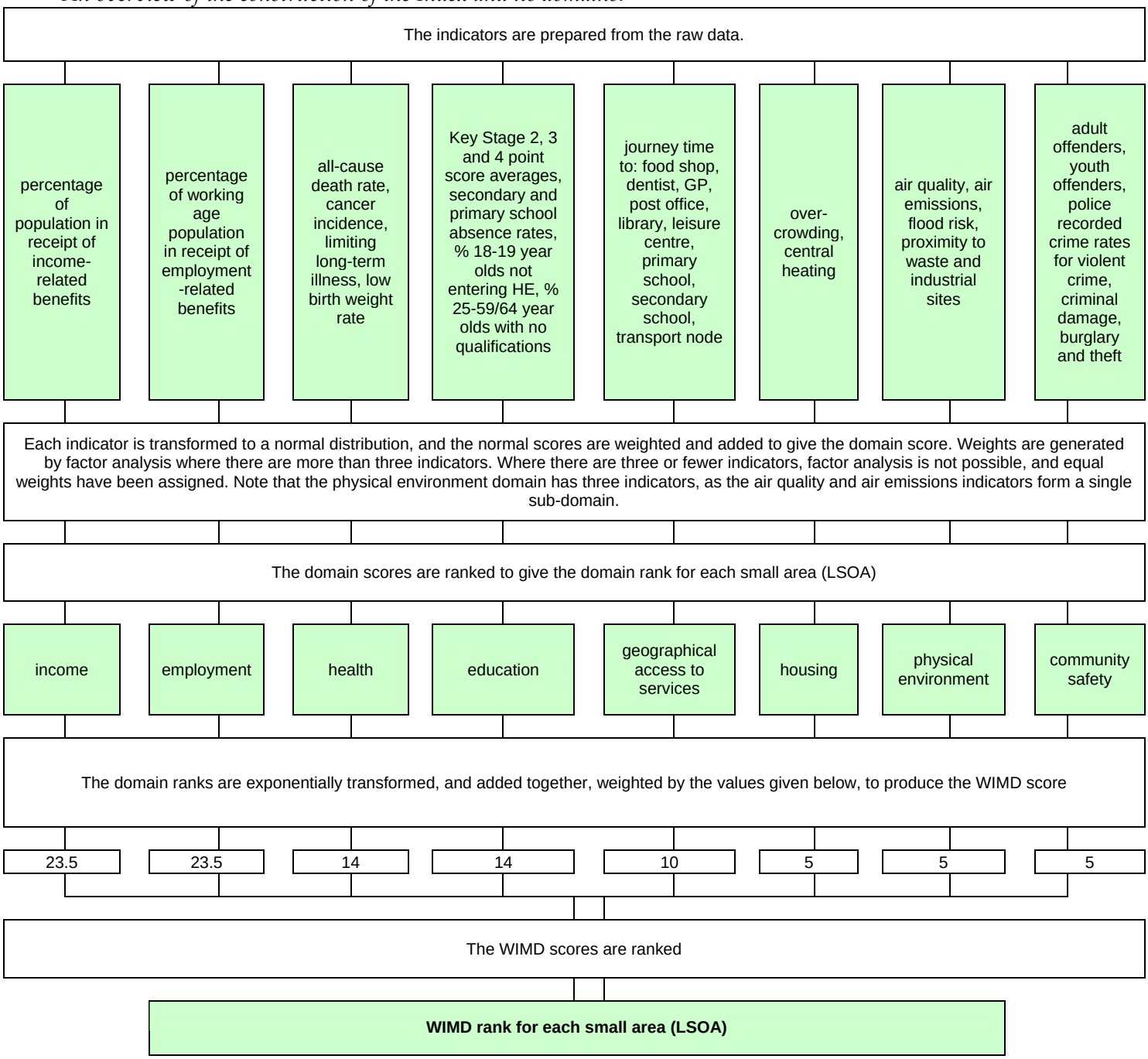
2 How the Index is constructed

The Index has three main components:

- the Index itself, which is a set of LSOA ranks;
- the LSOA ranks for each domain; and
- the underlying indicators, which are directly measurable, and combined to create the LSOA domain ranks. The units depend on what is being measured.

All of these components are calculated for each of the 1896 small areas (LSOAs) of Wales, and are published on the Welsh Government's StatsWales web pages. An overview of the construction of the Index is given in the diagram below.

An overview of the construction of the Index and its domains.



The Index is constructed from a weighted sum of the deprivation score for each domain. The weights reflect the importance of the domain as an aspect of deprivation, and the quality of the indicators available for that domain. The domains and their weights for WIMD 2011 are:

income	23.5%	geographical access to services	10%
employment	23.5%	community safety	5%
health	14%	physical environment	5%
education	14%	housing	5%

For WIMD 2011: Child Index the employment domain is dropped and its weight has been spread around the other domains:

income	35.3%	geographical access to services	12.2%
health	17.1%	community safety	6.1%
education	17.1%	physical environment	6.1%
Housing	6.1%		

A detailed analysis of the weighting system can be found in Paper 6 presented to the Steering Group on 22 October 2010. The weights used are decided by consensus based on extensive research.

The domains are in turn built up from sets of indicators, which are the measurable quantities which capture the concept of deprivation for each domain, e.g. the percentage of working age people in receipt of employment-based benefits for the employment domain; Key Stage scores in the education domain; crime rates in the community safety domain, etc. Indicators must be robust at the small area level and consistent across Wales and they are not all measured in the same units. In practice, this means that the Index is based largely on administrative data, with a limited number of Census variables where appropriate administrative data are not available.

Weightings within domains are derived by using factor analysis where appropriate. A detailed explanation of the factor analysis process is given in Annex C. Domains can be validated, either because they directly measure the factor (domain name) itself, for example, in the case of means-tested benefit, reliance on unemployment benefit, or by making comparisons with other research studies.

3 Geography: Lower Layer Super Output Areas

The Office for National Statistics developed a new geographic hierarchy called Super Output Areas (SOAs). They were designed to improve the reporting of small area statistics in England and Wales. Their first statistical application was for the Index of Deprivation for England in 2004, which led to them being widely used within local government. They have been increasingly used for data on the Neighbourhood Statistics (NeSS) website. It is anticipated that they will eventually become a standard for the production of National Statistics and will be used more generally.

Electoral divisions (previously known as wards) were the basic geographical units used for the Welsh Index of Multiple Deprivation 2000, but these vary greatly in size, from around 1,000 people to 20,000 (in Wales). This is not ideal for making comparisons throughout Wales, and it also means that data which can safely be released for larger electoral divisions may not be released for smaller ones due to disclosure rules (that is, the need to protect the confidentiality of individuals).

In addition, the boundaries of electoral divisions change. This creates problems when trying to compare data from different time periods. ONS developed a range of areas that would be of consistent size and whose boundaries would not change. These would be built from groups of the Output Areas (OAs) used for the 2001 Census, and would be known as Super Output Areas (SOAs).

There are three layers of SOAs: Lower Layer, Middle Layer, and Upper Layer. This was because disclosure requirements mean that some sets of data could be released for much smaller areas than others. To support a range of potential data requirements it was decided to create these three SOA layers. These are the constraints:

- A Lower Layer SOA must have a minimum population of 1,000. The mean size of all the Lower Layer SOAs must be close to 1,500. They are built from groups of Census OAs (usually between four and six).
- A Middle Layer SOA must have a minimum population of 5,000. The mean size of all the Middle Layer SOAs must be close to 7,200
- An Upper Layer SOA must have a minimum population of about 25,000: the mean population must be around 32,000.

There are 1896 Lower Layer SOAs in Wales (34,378 in England and Wales). They were generated by a computer program which merged OAs taking into account population size, mutual proximity, and social homogeneity. The boundaries were released in February 2004.

The following table gives the number of Lower Layer SOAs in each local authority in Wales.
Number of LSOAs by local authority

	number of Lower Layer SOAs		number of Lower Layer SOAs
Isle of Anglesey	44	Neath Port Talbot	91
Gwynedd	75	Bridgend	85
Conwy	71	The Vale of Glamorgan	78
Denbighshire	58	Cardiff	203
Flintshire	92	Rhondda, Cynon, Taff	152
Wrexham	85	Merthyr Tydfil	36
Powys	80	Caerphilly	110
Ceredigion	47	Blaenau Gwent	47
Pembrokeshire	71	Torfaen	60
Carmarthenshire	112	Monmouthshire	58
Swansea	147	Newport	94

There are 413 Middle Layer SOAs Wales (7,193 in England and Wales). They were generated in 2004, in two stages: a draft set was generated automatically (like the Lower Layer SOAs), then local authorities and other local agencies were invited to propose changes to the draft boundaries in order to establish areas that better met local needs.

Upper Super Output Areas (USOAs) in Wales were developed in 2008 by the Local Government Data Unit (LGDU) in association with the Welsh Local Government Association (WLGA). There are 94 USOAs in Wales with a mean population of around 32,000.

4 Income Domain

The purpose of this domain is to capture the extent of deprivation relating to income at a small area level across Wales. It focuses on the proportion of people living in households with income below a defined threshold or claiming benefits relating to low incomes. A full list of indicators and information about the indicators included in this domain follows. This includes indicators for both WIMD 2011 and WIMD 2011: Child Index.

Domain construction
Indicators were counts of unique individuals, in other words, duplicates were removed, so that indicators could simply be summed and expressed as a percentage of the total population of the LSOA. Each LSOA was then ranked in order, with the most deprived LSOA (rank 1) having the highest proportion of its population deprived and the least deprived LSOA (rank 1896) having the lowest proportion.

As with all domains, the final domain ranks were exponentially transformed, to form domain scores (see Appendix D). These domain scores were used in the calculation of the overall WIMD 2011 and WIMD 2011: Child Index ranks.

Methodological changes
The Employment and Support Allowance (ESA) has replaced Incapacity Benefit for new claimants since October 2008. The Income Support element of Incapacity Benefit is replaced with the income-based element of ESA.

The income domain has a weighting of 23.5% in the overall WIMD 2011 and a 35.3% weighting in WIMD 2011: Child Index.

Publication	WIMD 2011
Indicator	Income related benefits
Time Period	2009/10
Source	(1) The Department for Work and Pensions (DWP), Work and Pensions Longitudinal Study (WPLS) (2) Her Majesty's Revenue & Customs (3) Home office
Type	Percentage
Denominator	Population
Data availability	StatsWales
Additional Notes	*(1)Percentage in receipt of income related benefits* This indicator is formed by combined count of the yearly (Feb, May, Aug, Nov) average of Income Support (IS) claimants, Jobseekers Allowance (JSA) claimants, Pension Credit (PC) claimants, Income Based Employment and Support Allowance (ESA) claimants and the number of dependents on claimants of IS, JSA, PC, ESA, all divided by the total population. Data is collected from the WPLS is not subject to sampling error, as it is a 100 per cent survey. Income Support is intended to help people on low incomes who do not have to be available for employment. It can normally be claimed by people who are: of state pension age; working less than 16 hours a week (and/or with a partner working less than 24 hours a week); not required to be available for full-time employment; and in receipt of insufficient income to meet prescribed needs. The main beneficiary groups are: lone parents, on parental or paternity leave, carers, and refugees learning English. Pension credit is intended to help people on low incomes who do not have to be available for employment. It can normally be claimed by people who are: of state pension age; working less than 16 hours a week (and/or with a partner working less than 24 hours a week); not required to be available for full-time employment; and in receipt of insufficient income to meet prescribed needs. The main beneficiary groups are the long- and short-term sick and people with disabilities. JSA can be claimed by people aged 16 and over who are available for and actively seeking employment, including those in remunerative work for less than 16 hours a week on average, and by people on a government training scheme. All claimants of Job-Seeker's Allowance (JSA) are included in the claimant count, not just those in receipt of the income-based part. This is because it is not, at present, possible to separate the two components robustly at the small area level. This is not a methodological change from WIMD 2008, as all JSA claimants were included in the income domain in that edition too.

	(2)Dependent child receiving tax credits with income less than 60 per cent of the Wales median This is the addition of people claiming the Child Tax Credit and the Adult Tax credit. Child Tax Credit – Number of children in Welsh tax credits families (not in receipt of IS/JSA) with equivalised income less than 60% median income Working Tax Credit – an income-based benefit for working adults available to households with adults in one of four categories: 1) Work 16 hours or more a week and meet one of the following criteria: a) have responsibility for a child; b) have a disability that puts you at a disadvantage in getting a job; c) qualify for a 50-plus element; or 2) work 30 hours or more a week and are 25 or over. The methodology restricts inclusion to those below 60 per cent of the national median income. Tax credits provide an indication of families with low/moderate household incomes who are not on benefits. *(3) NASS-supported Asylum Seekers* The number of National Asylum Support Service (NASS) supported Asylum seekers at the end of December 2010 are also added to this indicator. This indicator represents the presence of asylum seekers in Wales and contains people who are not eligible for the other benefits in the domain, yet are income deprived.
Publication	WIMD 2011: Child Index
Indicator	Income related benefits – Child Index
Time Period	2009/10
Source	(1) The Department for Work and Pensions (DWP), Work and Pensions Longitudinal Study (WPLS). (2) Her Majesty's Revenue & Customs
Type	Percentage
Denominator	LSOA Population, aged 0-18 years old
Data availability	StatWales
Additional Notes	*(1)Children in households receiving income related benefits* This indicator is formed by combined count of the yearly (Feb, May, Aug, Nov) average of Income Support (IS) claimants, Jobseekers Allowance (JSA) claimants, Pension Credit (PC) claimants, Income Based Employment and Support Allowance (ESA) claimants and the number of dependents on claimants of IS, JSA, PC, ESA, all divided by the total population. Data is collected from the WPLS is not subject to sampling error, as it is a 100 per cent survey. Income Support is intended to help people on low incomes who do not have to be available for employment. It can normally be claimed by people who are: of state pension age; working less than 16 hours a week (and/or with a partner

working less than 24 hours a week); not required to be available for full-time employment; and in receipt of insufficient income to meet prescribed needs. The main beneficiary groups are: lone parents, on parental or paternity leave, carers, and refugees learning English.

Pension credit is intended to help people on low incomes who do not have to be available for employment. It can normally be claimed by people who are: of state pension age; working less than 16 hours a week (and/or with a partner working less than 24 hours a week); not required to be available for full-time employment; and in receipt of insufficient income to meet prescribed needs. The main beneficiary groups are the long- and short-term sick and people with disabilities.

JSA can be claimed by people aged 16 and over who are available for and actively seeking employment, including those in remunerative work for less than 16 hours a week on average, and by people on a government training scheme.

All claimants of Job-Seeker's Allowance (JSA) are included in the claimant count, not just those in receipt of the income-based part. This is because it is not, at present, possible to separate the two components robustly at the small area level. This is not a methodological change from WIMD 2008, as all JSA claimants were included in the income domain in that edition too.

(2)Children in households receiving tax credits with income less than 60 per cent of the Wales median
This is the addition of people claiming the Child Tax Credit and the Adult Tax credit.

Child Tax Credit – Number of children in Welsh tax credits families (not in receipt of IS/JSA) with equivalised income less than 60% median income

Working Tax Credit – an income-based benefit for working adults available to households with adults in one of four categories:
1) Work 16 hours or more a week and meet one of the following criteria:
    d) have responsibility for a child;
    e) have a disability that puts you at a disadvantage in getting a job;
    f) qualify for a 50-plus element;
or
2) work 30 hours or more a week and are 25 or over.

The methodology restricts inclusion to those below 60 per cent of the national median income.

Tax credits are they provide an indication of families with low/moderate household incomes who are not on benefits.

Employment Domain

The purpose of this domain is to capture the extent of deprivation relating to employment. Note that there is no employment domain in WIMD 2011: Child Index, so indicators listed only form part of WIMD 2011.

Domain construction

Indicators were counts of unique individuals, duplicates were removed, so that indicators could simply be summed and expressed as a percentage of the working age population of the LSOA. Data was received from DWP as percentages rounded to the nearest whole number.

Each LSOA was then ranked in order, with the most deprived LSOA (rank 1) having the highest proportion of its population deprived and the least deprived LSOA (rank 1896), having the lowest proportion. As with all domains, the final domain ranks were exponentially transformed, to form domain scores (see Appendix D). These domain scores were used in the calculation of the overall WIMD 2011.

Methodological changes

The Employment and Support Allowance (ESA) has replaced Incapacity Benefit for new claimants since October 2008.

For WIMD 2011, data from DWP has been received as percentages by LSOA rounded to the nearest whole number. This means that it is possible for several LSOAs to have the same indicator values (percentage of working age people in receipt of employment-related benefits) and therefore there are a relatively large number of tied ranks in this domain. This means that the standard rank ranges (1-190, 191-380 etc) do not correspond exactly to the usual 10, 20, 30 per cent etc. of the number of LSOAs in Wales.

The employment domain has a weighting of 23.5% in the overall WIMD 2011.

Indicators

Publication	WIMD 2011 only
Indicator	In receipt of employment related benefits
Time Period	2009/10
Source	The Department for Work and Pensions (DWP), Work and Pensions Longitudinal Study (WPLS) which is a 100 per cent data source that is not subject to any sampling error
Type	Percentage
Denominator	Working age population
Data availability	StatsWales
Additional Notes	The yearly average comprises of data from November 2009, February 2010, May 2010 and August 2010. This was formed from the combined count of : 1. Claimants of Incapacity Benefit/ Severe Disablement Allowance (IBSDA) Incapacity Benefit claimants are those of working age who are unable to work due to illness or disability and who meet criteria contribution conditions. All three rate categories of claimants are included (short-term lower, short-term higher and long term). Severe disablement Allowance is available to those people who are incapable of work and do not satisfy the contribution conditions of Incapacity Benefit. 2. Claimants of New Deal for Young People (NDYP), New Deal for 25+ (ND25+), and New Deal for Lone Parents (NDLP) This indicator captures young people (18-24) who have been claiming JSA for at least 6 months and persons 25+ who have been claiming JSA for at least 2 years and have therefore moved into the New Deal programme. This indicator was chosen as it captures people who are on a scheme which encourages them back to work, but who have not found employment yet and are not picked up on the JSA count. The New Deal for Lone Parents (NDLP) is designed specifically to help lone parents aged 16 or over who are not working, or who are working less than 16 hours per week, and who have sole responsibility for the care of a child(ren) under the age of 16. 3. Claimants of Jobseekers Allowance (JSA) This indicator captures all persons of working age (18-59 for women and 18-64 for men), who are claiming Job Seekers Allowance, available to those who are actively seeking work, but not in work. 4. Claimants of Employment and Support Allowance (ESA) This indicator captures all persons of working age (18-59 for women and 18-64 for men), who are claiming Employment and Support Allowance. It is available to people who are unable to work because of illness or disability, and those who are able to work. ESA has replaced Incapacity Benefit for new claimants since October 2008.

Health Domain

The purpose of this domain is to capture the extent of deprivation relating to good health. This includes indicators for both WIMD 2011 and WIMD 2011: Child Index.

Domain construction

Indicators were combined using factor analysis, to determine how much weighting each indicator should have within the domain (see Appendix C). The resultant weights for each indicator were as follows for WIMD 2011:

0.38 Limiting long-term illness
0.39 All cause death rate
0.13 Cancer incidence
0.10 Low Birth Weight

For WIMD 2011: Child Index, there are only two indicators, which have an equal weighting:

0.50 Percentage of live single births <2.5kg
0.50 Limiting long-term illness in children

Each LSOA was then ranked in order, with the most deprived LSOA (rank 1) having the highest proportion of its population deprived and the least deprived LSOA (rank 1896) having the lowest proportion. As with all domains, the final domain ranks were exponentially transformed, to form domain scores (see Appendix D). These domain scores were used in the calculation of the overall WIMD 2011 and WIMD 2011: Child Index ranks.

Methodological changes

Following the discovery of some anomalous data points in the health domain of WIMD 2008, the WIMD Steering Group approved a change to the age-sex standardisation methodology of the health indicators from direct to indirect standardisation for WIMD 2011. Indirect standardisation is a more robust approach for small number statistics such as WIMD indicators. The change of methodology limited effects and is greatest where use of direct standardisation for very small numbers produced anomalous results for WIMD 2008. The new methodology corrects this problem.

Full analysis of the effects of direct versus indirect standardisation is presented in the paper and minutes of the Health Domain Working Group in May 2011. The discussion and agreement of the WIMD Steering Group can be found in the minutes of the Steering Group meeting of 2 June 2011

The health domain has a 14% weighting in WIMD 2011 and a 17.1% weighting in the WIMD 2011: Child Index.

Indicators

Publication	WIMD 2011 only
Indicator	Cancer incidence
Time Period	2000-2009
Source	Velindre NHS Trust
Type	Number per 100,000
Denominator	Indirect age-sex standardised populations
Data availability	StatsWales
Additional Notes	Count of all cases of cancer includes all malignancies excluding non melanoma skin cancer. Cancer incidence data are robust both in terms of the long standing central collection of the data and the numbers of cases involved even at a small area level.

Publication	WIMD 2011 only
Indicator	All-cause death rate
Time Period	2001-2009
Source	Office of National Statistics (ONS)
Type	Number per 100,000
Denominator	Indirect age-sex standardised populations
Data availability	StatsWales
Additional Notes	Poor health manifests itself both through a poorer quality of live but also in lower life expectancy which can be captured through age and sex standardised death rates.

Publication	WIMD 2011 only
Indicator	Limiting long-term illness
Time Period	2001
Source	Census, Office of National Statistics (ONS)
Type	Number per 100,000
Denominator	Indirect age-sex standardised populations
Data availability	StatsWales
Additional Notes	A limiting long-term illness covers any long-term illness, health problem or disability that limits daily activities or work. The question in the Census on limiting long-term illness is well established and has been shown by numerous studies to be a valid measure of morbidity in the community.

Publication	WIMD 2011: Child Index only
Indicator	Limiting long-term illness, in Children
Time Period	2001
Source	Census, Office of National Statistics (ONS)
Type	Number per 100,000
Denominator	LSOA population, 0 - 17 year olds only
Data availability	StatsWales
Additional Notes	A limiting long-term illness covers any long-term illness, health problem or disability that limits daily activities or work. The question in the Census on limiting long-term illness is well established and has been shown by numerous studies to be a valid measure of morbidity in the community.

Publication	WIMD 2011 and WIMD 2011: Child Index
Indicator	Low weight single births
Time Period	2001-2009
Source	Office of National Statistics (ONS)
Type	Percentage
Denominator	Number of live single births.
Data availability	StatsWales
Additional Notes	Low birth weight is defined as birth weight less than 2500g. Evidence suggests that low birth rate is linked to the mother's lifestyle and health. Low birth rate can also cause problems for a baby in later life increasing the risk of chronic diseases.

Education Domain

The purpose of this domain is to capture the extent of deprivation relating to education, training and skills. It is designed to reflect educational disadvantage within an area, by capturing low attainment among children and young people and the lack of qualifications and skills in adults. This includes indicators for both WIMD 2011 and WIMD 2011: Child Index.

Domain construction
Indicators were combined using factor analysis, to determine how much weighting each indicator should have within the domain (see Appendix C). The resultant weights for each indicator were as follows for WIMD 2011:

0.08 Key Stage 2 average point scores
0.22 Key Stage 3 average point scores
0.10 Key Stage 4 wider average point scores
0.18 Percentage of people not entering higher education age 18-19
0.20 Percentage of adults aged 25-59/64 with no qualifications
0.10 Percentage primary school half day absence
0.13 Percentage secondary school half day absence

The weighting for the five indicators in WIMD 2011: Child Index were as follows:

0.11 Key Stage 2 average point scores
0.33 Key Stage 3 average point scores
0.17 Key Stage 4 wider average point scores
0.17 Percentage primary school half day absence
0.22 Percentage secondary school half day absence

Each LSOA was then ranked in order, with the most deprived LSOA (rank 1) having the highest proportion of its population deprived and the least deprived LSOA (rank 1896) having the lowest proportion. As with all domains, the final domain ranks were exponentially transformed, to form domain scores (see Appendix D). These domain scores were used in the calculation of the overall WIMD 2011 and WIMD 2011: Child Index ranks.

Methodological changes
Pupil postcodes are now available for both primary and secondary absences. This is an improvement over WIMD 2008, where the school-level data had to be apportioned to postcodes.

The Key Stage 4 (GCSE level) point scoring system changed in 2008, and it also now includes qualifications other than GCSEs. Data are no longer available for the old point system. The inclusion of qualifications other than GCSEs improves this indicator, but both the level and the ratio between successive grades are changed, which may lead to differences in the ranking.

The education domain has a 14% weighting in WIMD 2011 and a 17.1% weighting in the WIMD 2011: Child Index

Indicators

Publication	WIMD 2011 and WIMD 2011: Child Index
Indicator	Key Stage 2 average point scores
Time Period	2008-10
Source	National Pupil Database, Pupil Level Annual School Census (PLASC) and National Data Collection.
Type	Point score
Denominator	Number of Key Stage 2 pupils.
Data availability	StatsWales
Additional Notes	Average points scores pupils are assessed by teachers in Year 6 (final year of primary school) for 2008, 2009 and 2010. Data from PLASC is able to be matched to LSOAs, using a postcode to LSOA look-up provided by the Geography and Technology department of the Welsh Government. A 3-year average is used to reduce the impact of having small numbers of pupils at LSOA level. Because not all children are assessed in Welsh as a first language at key stages 2 and 3, the highest score in English and Welsh was taken along with the score in Mathematics and Science to provide comparability across Wales.

Publication	WIMD 2011 and WIMD 2011: Child Index
Indicator	Key Stage 3 average point scores
Time Period	2008-10
Source	National Pupil Database, Pupil Level Annual School Census (PLASC) and National Data Collection.
Type	Point score
Denominator	Number of Key Stage 3 pupils.
Data availability	StatsWales
Additional Notes	Average points scores pupils are assessed by teachers in Year 9 for 2008, 2009 and 2010. Data from PLASC is able to be matched to LSOAs, using a postcode to LSOA look-up provided by the Geography and Technology department of the Welsh Government. A 3-year average is used to reduce the impact of having small numbers of pupils at LSOA level. Because not all children are assessed in Welsh as a first language at key stages 2 and 3, the highest score in English and Welsh was taken along with the score in Mathematics and Science to provide comparability across Wales.

Publication	WIMD 2011 and WIMD 2011: Child Index
Indicator	Key Stage 4 wider average point scores
Time Period	2009-10
Source	National Pupil Database, Pupil Level Annual School Census (PLASC) and Welsh Examinations Database.
Type	Point score
Denominator	Number of Key Stage 4 pupils.
Data availability	StatsWales
Additional Notes	The Key Stage 4 (GCSE level) point scoring system changed in 2008 and now includes other equivalent qualifications that are not GCSEs. Data are not available for the old point scoring system, so direct comparison with WIMD 2008 indicator data is not possible. The inclusion of different qualifications improves this indicator, but both the level and ratio between successive grades are changed, which may lead to differences in ranking, when compared to WIMD 2008 ranks.

Publication	WIMD 2011 and WIMD 2011: Child Index
Indicator	Percentage of people not entering higher education aged 18-19
Time Period	1998-2005
Source	Higher Education Funding Council for England (HEFCE)
Type	Percentage
Denominator	Number of 18-19 year olds
Data availability	StatsWales
Additional Notes	This indicator has not been updated since WIMD 2008 was published as HEFCE were unable to supply the relevant data.

Publication	WIMD 2011 and WIMD 2011: Child Index
Indicator	Percentage of adults aged 25 to 59 or 64 with no qualifications.
Time Period	2001
Source	ONS Census
Type	Percentage
Denominator	Number of adults aged 25 to 59 or 64 year olds.
Data availability	StatsWales
Additional Notes	The WIMD 2008 indicator has not been updated as data from the 2011 Census was not available. The indicator refers to adults aged 25 to retirement age, which at the time the indicator was produced, was 60 for females and 65 for males. This indicator includes those who hold no qualifications or qualifications which do not reach the standard to be categorised as Level 1 qualifications. These persons were chosen because persons with no qualifications are less likely to have functional literacy and numeracy skills, more likely to be economically inactive and on average earn less.

Publication	WIMD 2011 and WIMD 2011: Child Index
Indicator	Percentage of primary school half day absence
Time Period	2008/09-2009/10
Source	Pupil Level Annual School Census (PLASC) and Pupil Attendance records.
Type	Percentage
Denominator	Number of half day school sessions
Data availability	StatsWales
Additional Notes	Data from PLASC is able to be matched to LSOAs, using a postcode to LSOA look-up provided by the Geography and Technology department of the Welsh Government. Data on the number of school sessions missed due to authorised and unauthorised absence is collected for the entire academic year. Independent schools were not included due to unavailability of data for all schools. Pupils from Welsh LSOAs that go to schools in England we reincluded in this.

Publication	WIMD 2011 and WIMD 2011: Child Index
Indicator	Percentage of secondary school half day absence
Time Period	2008/09-2009/10
Source	Pupil Level Annual School Census (PLASC) and Pupil Attendance records.
Type	Percentage
Denominator	Number of half day school sessions
Data availability	StatsWales
Additional Notes	Data on the number of school sessions missed due to authorised and unauthorised absence is collected from the start of the academic year up to the date of the late May bank holiday. Independent schools were not included due to unavailability of data for all schools. Pupils from Welsh LSOAs that go to schools in England we included in this. Postcodes were supplied with the data, which were then allocated to LSOAs

Community Safety

The purpose of this domain is to capture the extent of deprivation relating to living in a safe community. Safety includes levels of household and personal crime and quality of experience in public places compatible with access to ordinary work, leisure and social relationships. For this domain the indicators for both WIMD 2011 and WIMD 2011: Child Index are the same.

Domain construction
Indicators were combined using factor analysis, to determine how much weighting each indicator should have within the domain (see Appendix C). The resultant weights for each indicator were as follows for WIMD 2011 and WIMD 2011: Child Index:

0.20 Percentage of adult offenders
0.06 Police recorded burglary
0.35 Police recorded criminal damage
0.04 Fire incidence
0.07 Police recorded theft
0.19 Police recorded violent crime
0.08 Percentage of youth offenders

Each LSOA was then ranked in order, with the most deprived LSOA (rank 1) having the highest proportion of its population deprived and the least deprived LSOA (rank 1896) having the lowest proportion. As with all domains, the final domain ranks were exponentially transformed, to form domain scores (see Appendix D). These domain scores were used in the calculation of the overall WIMD 2011 and WIMD 2011: Child Index ranks.

Methodological changes
In WIMD 2011 and WIMD 2011: Child Index, the violent crime rate is based on a single year (2009/10) of data, instead of the average over two years of data. This is due to a crime coding issue in the 2008/9 data year.

The community safety domain has a 5% weighting in WIMD 2011 and a 6.1% weighting in the WIMD 2011: Child Index

Indicators

Publication	WIMD 2011 and WIMD 2011: Child Index
Indicator	Percentage of adult offenders
Time Period	2008/9 – 2009/10
Source	Wales Probation Trust
Type	Percentage
Denominator	Resident population aged 18+, less prison population aged 18+
Data availability	StatsWales
Additional Notes	Adult offenders are often drawn from the most deprived section of society. Offender information complements crime records and was collected by the probation service. Within Wales there are four probation service areas (North Wales, Dyfed Powys, South Wales and Gwent). The offenders dealt with by those teams are aged 18 and over. Where a person had more than one address during the period each address was counted provided it was in a different LSOA. Where there were multiple offenders at a single address a weight of 2 was applied. Offenders residing in institutions were included in the count if a residential address was provided.

Publication	WIMD 2011 and WIMD 2011: Child Index
Indicator	Police recorded burglary
Time Period	2008/9 – 2009/10
Source	Welsh Police Forces
Type	Percentage
Denominator	Number of dwellings, plus total number of business addresses
Data availability	StatsWales
Additional Notes	The incidents (crimes) were recorded by the police forces in Wales (North Wales, Dyfed Powys, South Wales and Gwent). The incidents were located to the point at which they occurred and allocated to the appropriate lower super output area (LSOA). They were of selected types which affect individuals or businesses; violence, burglary, criminal damage and theft. The criteria for selecting incidents by crime code is similar to that used in England, but a small number of additional codes were included. The additional codes produced only a small number of additional incidents.

Publication	WIMD 2011 and WIMD 2011: Child Index
Indicator	Police recorded criminal damage
Time Period	2008/9 – 2009/10
Source	Welsh Police Forces
Type	Percentage
Denominator	Resident population excluding prisoners, plus total non-resident workplace population
Data availability	StatsWales
Additional Notes	The incidents (crimes) were recorded by the police forces in Wales (North Wales, Dyfed Powys, South Wales and Gwent). The incidents were located to the point at which they occurred and allocated to the appropriate lower super output area (LSOA). They were of selected types which affect individuals or businesses; violence, burglary, criminal damage and theft. The criteria for selecting incidents by crime code is similar to that used in England, but a small number of additional codes were included. The additional codes produced only a small number of additional incidents.

Publication	WIMD 2011 and WIMD 2011: Child Index
Indicator	Fire Incidence
Time Period	2009/10 – 2010/11
Source	Welsh Government
Type	Percentage
Denominator	Resident population
Data availability	StatsWales
Additional Notes	Incidents of primary fires with the addition of “derelict vehicle” fires were collected as counts by LSOA. Secondary "derelict vehicle" fires were only included if available at the LSOA level. Incidents requiring call out of fire and rescue services are related to deprivation and more likely within disadvantaged groups. Primary fires include “all fires in buildings, vehicles and outdoor structures or any fire involving casualties, rescues, or fires attended by five or more appliances”. Secondary fires are “the majority of outdoor fires including grassland and refuse fires unless they involve casualties or rescues, property loss or five or more appliances attend”. Primary fires, which relate better to property and people, with the addition of “derelict vehicle” fires that tend to occur in deprived areas, was used. The secondary "derelict vehicle" fires were only included if available at the LSOA level.

Publication	WIMD 2011 and WIMD 2011: Child Index
Indicator	Police recorded theft
Time Period	2008/9 – 2009/10
Source	Welsh Police Forces
Type	Percentage
Denominator	Resident population excluding prisoners, plus total non-resident workplace population.
Data availability	StatsWales
Additional Notes	The incidents (crimes) were recorded by the police forces in Wales (North Wales, Dyfed Powys, South Wales and Gwent). The incidents were located to the point at which they occurred and allocated to the appropriate lower super output area (LSOA). They were of selected types which affect individuals or businesses; violence, burglary, criminal damage and theft. The criteria for selecting incidents by crime code is similar to that used in England, but a small number of additional codes were included. The additional codes produced only a small number of additional incidents.

Publication	WIMD 2011 and WIMD 2011: Child Index
Indicator	Police recorded violent crime
Time Period	2009/10
Source	Welsh Police Forces
Type	Percentage
Denominator	Resident population excluding prisoners, plus total non-resident workplace population.
Data availability	StatsWales
Additional Notes	In WIMD 2011 and the Child Index 2011 the violent crime indicator is based on a single year of data, instead of an average over 2 years of data. This is due to a crime coding issue in the 2008/9 data. The incidents (crimes) were recorded by the police forces in Wales (North Wales, Dyfed Powys, South Wales and Gwent). The incidents were located to the point at which they occurred and allocated to the appropriate lower super output area (LSOA). They were of selected types which affect individuals or businesses; violence, burglary, criminal damage and theft. The criteria for selecting incidents by crime code is similar to that used in England, but a small number of additional codes were included. The additional codes produced only a small number of additional incidents.

Publication	WIMD 2011 and WIMD 2011: Child Index
Indicator	Percentage of Youth offenders
Time Period	2008/9 – 2009/10
Source	Ministry of Justice
Type	Percentage
Denominator	Resident population aged 10-17, less prison population age 10-17
Data availability	StatsWales
Additional Notes	Offenders often come from the most deprived section of society. Offenders at liberty is a measure to indicate issues relating to young people that indicate their need and the social problems they cause. This indicator complements the recorded crime data by including the offenders at liberty in an area. Within Wales there are 18 Youth Offending Teams. The offenders dealt with by those teams are aged from 10 to 17 years of age. Where a person had more than one address during the period each addresses in was counted provided it was in a different LSOA. Where there were multiple offenders at a single address a weight of 2 was applied. Care homes and institutions were excluded from the counts.

Geographical Access to Services Domain

The purpose of this domain is to illustrate the deprivation as a result of a household's inability to access a range of services, considered necessary for day-to-day living. The indicators are measured by the time taken to reach the service, using a bus, walking or a combination of the two. The geographical access to services domain was not updated for WIMD 2011 or WIMD 2011: Child Index. Therefore, it is exactly the same as the domain in WIMD 2008 and WIMD 2008: Child Index. A full list of indicators and information about the indicators included in this domain follows.

Domain construction
Indicators were combined using factor analysis, to determine how much weighting each indicator should have within the domain (see Appendix C). The resultant weights for each indicator were as follows for WIMD 2011:

0.12 Mean bus + walking journey time to NHS dentist
0.14 Mean bus + walking journey time to food shop
0.18 Mean bus + walking journey time to GP
0.06 Mean bus + walking journey time to leisure centre
0.09 Mean bus + walking journey time to library
0.10 Mean bus + walking journey time to Post Office
0.16 Mean bus + walking journey time to primary school
0.07 Mean bus + walking journey time to secondary school
0.07 Mean bus + walking journey time to transport node

The weighting for the four indicators in WIMD 2011: Child Index was as follows:

0.40 Mean bus + walking journey time to leisure centre
0.17 Mean bus + walking journey time to library
0.17 Mean bus + walking journey time to primary school
0.26 Mean bus + walking journey time to secondary school

Each LSOA was then ranked in order, with the most deprived LSOA (rank 1) having the highest proportion of its population deprived and the least deprived LSOA (rank 1896) having the lowest proportion. As with all domains, the final domain ranks were exponentially transformed, to form domain scores (see Appendix D). These domain scores were used in the calculation of the overall WIMD 2011 and WIMD 2011: Child Index ranks.

Methodological changes
There are no methodological changes as the domain has not been updated since WIMD 2008 because of data availability issues.

Calculating the Access to Services Domain
Grid reference of service point locations were provided by the Geography and Technology Unit of the Welsh Government Knowledge and Analytical Services Unit. Royal Mail address points were used as the origin points for journeys that individuals would need to make to these service points. Commercial address points were excluded in urban areas but included in rural areas to ensure farms were not excluded (not all of which have farm in the description in the address point file).

Calculations were undertaken in a package developed for local transport planning. The road network in Wales as well as bus routes and timetables were input to the package. The time needed to get from each origin to each service point was then calculated based on the distance and availability of buses. A maximum walk of 800 metres was set for the start and end parts of the journey, i.e. from home to the bus stop and from the bus stop to the service, or direct to the service point if that should apply.

Up to 10 of the shortest trips were recorded from each household to each service type – in excess of 10 million trips were recorded.

The indicators were then calculated as the time taken for every household in a LSOA to travel to each of the services identified in the domain using the average of the 10 shortest trips.

Indicators

Publication	WIMD 2011 only
Indicator	Mean bus and walking journey time to NHS dentist
Time Period	2008
Source	Public Transport and Road Network Postal Address File, Ordnance survey.
Type	Time in minutes.
Denominator	Number of journeys.
Data availability	StatsWales
Additional Notes	The average time taken for every household in a LSOA to travel to a NHS dentist by walking or using public bus services or both. This indicator is based on all dentists offering NHS treatment. It looks solely at an individual's ability to access a surgery and does not take into account whether spaces are available, i.e. the indicator is measuring purely geographical access.

Publication	WIMD 2011 only
Indicator	Mean bus and walking journey time to food shop
Time Period	2008
Source	Public Transport and Road Network Postal Address File, Ordnance survey.
Type	Time in minutes.
Denominator	Number of journeys.
Data availability	StatsWales
Additional Notes	The average time taken for every household in a LSOA to travel to a food shop by walking or using public bus services or both. This indicator is intended to cover the purchase of basic provisions (e.g. bread and milk). Service points include premises from the local corner shop up to large supermarkets.

Publication	WIMD 2011 only
Indicator	Mean bus and walking journey time to GP
Time Period	2008
Source	Public Transport and Road Network Postal Address File, Ordnance survey.
Type	Time in minutes.
Denominator	Number of journeys.
Data availability	StatsWales
Additional Notes	The average time taken for every household in a LSOA to travel to a GP surgery by walking or using public bus services or both. This indicator is intended to cover the day-to-day need for primary health care. This indicator includes all GP surgeries, although the services available across GP surgeries vary, the basic services are offered by all surgeries.

Publication	WIMD 2011 and WIMD 2011: Child Index
Indicator	Mean bus and walking journey time to leisure centre
Time Period	2008
Source	Public Transport and Road Network Postal Address File, Ordnance survey.
Type	Time in minutes.
Denominator	Number of journeys.
Data availability	StatsWales
Additional Notes	The average time taken for every household in a LSOA to travel to a leisure centre by walking or using public bus services or both. This indicator is intended to cover an individual's ability to access facilities important for health and well being.

Publication	WIMD 2011 and WIMD 2011: Child Index
Indicator	Mean bus and walking journey time to a public library
Time Period	2008
Source	Public Transport and Road Network Postal Address File, Ordnance survey.
Type	Time in minutes.
Denominator	Number of journeys.
Data availability	StatsWales
Additional Notes	The average time taken for every household in a LSOA to travel to a public library by walking or using public bus services or both. This indicator covers all static libraries and reflects a library's role as a vital modern communications service, with online access to further advice and information, rather than just access to the more traditional services (e.g. book lending). For WIMD 2008/2011 mobile libraries were also included.

Publication	WIMD 2011 only
Indicator	Mean bus and walking journey time to Post office
Time Period	2008
Source	Public Transport and Road Network Postal Address File, Ordnance survey.
Type	Time in minutes.
Denominator	Number of journeys.
Data availability	StatsWales
Additional Notes	The average time taken for every household in a LSOA to travel to a post office by walking or using public bus services or both. This indicator encapsulates accessibility to a post office. Post offices are used weekly by most of the community and are a vital communications source. This may include obtaining advice and assistance, sending and collection of mail/parcels, collection of benefits, payment of bills, withdrawal of money and many other services.

Publication	WIMD 2011 and WIMD 2011: Child Index
Indicator	Mean bus and walking journey time to primary school
Time Period	2008
Source	Public Transport and Road Network Postal Address File, Ordnance survey.
Type	Time in minutes.
Denominator	Number of journeys.
Data availability	StatsWales
Additional Notes	The average time taken for every household in a LSOA to travel to a primary school by walking or using public bus services or both. This indicator is designed to reflect the access to a primary school of children aged 4 to 11 to a primary school. Access is defined purely on the child's ability to access any primary school and takes no account of the school actually attended.

Publication	WIMD 2011 and WIMD 2011: Child Index
Indicator	Mean bus and walking journey time to secondary school
Time Period	2008
Source	Public Transport and Road Network Postal Address File, Ordnance survey.
Type	Time in minutes.
Denominator	Number of journeys.
Difference for WIMD 2011: Child Index?	None.
Data availability	StatsWales
Additional Notes	The average time taken for every household in a LSOA to travel to a secondary school by walking or using public bus services or both. This indicator is designed to reflect the access needs of children aged 11 to 16 to a secondary school. Access is defined purely on the child's ability to access any secondary school and takes no account of the school actually attended.

Publication	WIMD 2011 only
Indicator	Mean bus and walking journey time to transport node
Time Period	2008
Source	Public Transport and Road Network Postal Address File, Ordnance survey.
Type	Time in minutes.
Denominator	Number of journeys.
Data availability	StatsWales
Additional Notes	The average time taken for every household in a LSOA to travel to a transport node by walking or using public bus services or both. This is a new indicator for WIMD 2008 to show access to long distance transport services. The proximity of transport nodes to each household was computed for coaches and rail routes.

Housing Domain

The purpose of this domain is to capture deprivation through a lack of adequate housing. A full list of indicators and information about the indicators included in this domain follows. The housing domain was not updated for WIMD 2011 or WIMD 2011: Child Index. Therefore, it is exactly the same as the domain in WIMD 2008 and WIMD 2008: Child Index. The indicators for WIMD 2011: Child Index are the same as those used in WIMD 2011, except WIMD 2011 indicators refer to all people living in households and WIMD 2011: Child Index indicators refer to children aged 0-18 years old.

Domain construction
There are only two indicators in the housing domain, which have an equal weighting:

0.50 Lack of central heating
0.50 Overcrowding

Each LSOA was then ranked in order, with the most deprived LSOA (rank 1) having the highest proportion of its population deprived and the least deprived LSOA (rank 1896) having the lowest proportion. As with all domains, the final domain ranks were exponentially transformed, to form domain scores (see Appendix D). These domain scores were used in the calculation of the overall WIMD 2011 and WIMD 2011: Child Index ranks.

Methodological changes
There are no methodological changes as the domain has not been updated since WIMD 2008 because indicator data is sourced from the ten yearly Census, which has not been updated since WIMD 2008. There are no other suitable sources for the WIMD housing domain at present either.

Indicators

Publication	WIMD 2011
Indicator	Percentage of population living in households with no central heating.
Time Period	2001
Source	Census, Office of National Statistics.
Type	Percentage
Denominator	Number of people living in households.
Data availability	(StatsWales table)
Additional Notes	Indicator data is exactly the same as used in WIMD 2008 as there are not more recent data available.

Publication	WIMD 2011
Indicator	Percentage of population living in overcrowded households
Time Period	2001
Source	Census, ONS
Type	Percentage
Denominator	Number of people living in households, excluding all people living in student households.
Data availability	(StatsWales table)
Additional Notes	Where there are one or more rooms, too few for a household, the house will have an occupancy rating of -1 or less. This indicator counts all people living in a house with an occupancy rating of -1 or lower, based on the 2001 Census. This indicator data is exactly the same as used in WIMD 2008 as there are not more recent data available.

Publication	WIMD 2011: Child Index
Indicator	Percentage of population living in households with no central heating.
Time Period	2001
Source	Census, ONS.
Type	Percentage
Denominator	Population of children (0-18 years old) living in households, excluding all people living in student households.
Data availability	(StatsWales table)
Additional Notes	Indicator data is exactly the same as used in WIMD 2008 as there are not more recent data available.

Publication	WIMD 2011: Child Index
Indicator	Percentage of children living in overcrowded households
Time Period	2001
Source	Census, ONS
Type	Percentage
Denominator	Population of children (0-18 year olds) living in households, excluding all people living in student households.
Data availability	(StatsWales table)
Additional Notes	Where there are one or more rooms, too few for a household, the house will have an occupancy rating of -1 or less. This indicator counts all people living in a house with an occupancy rating of -1 or lower, based on the 2001 Census. This indicator data is exactly the same as used in WIMD 2008 as there are not more recent data available.

Physical Environment Domain

The purpose of this domain is to measure environmental factors that may impact on quality of life in an area. Environmental deprivation is generally not correlated with social or economic deprivation in Wales. (ref: Walker et al 2003). Data for Wales is not available on actual impacts on quality of life; so the indicators chosen have good data availability and indicate the potential for reduced quality of life. A full list of indicators and information about the indicators included in this domain follows. The indicators used in WIMD 2011: Child Index are the same as those used in WIMD 2011.

Domain Construction
This domain was broken down into 3 sub-domains of similar measures. Each sub-domain was given an equal (1/3) weighting. The air pollution sub-domain contains 2 indicators, (air quality and air emissions) which are equally weighted within that sub domain. This means that for the overall domain, air quality and air emissions have a 0.17 weight.

0.17 Air quality
0.17 Air emissions
0.33 Flood risk
0.33 Proximity to waste disposal and industrial sites

Each LSOA was then ranked in order, with the most deprived LSOA (rank 1) having the highest proportion of its population deprived and the least deprived LSOA (rank 1896) having the lowest proportion. As with all domains, the final domain ranks were exponentially transformed, to form domain scores (see Appendix D). These domain scores were used in the calculation of the overall WIMD 2011 and WIMD 2011: Child Index ranks.

Methodological Changes
The methodology used to create the indicators of the physical environment domain is the same for WIMD 2011 and WIMD 2011: Child Index is as it was for WIMD 2008. However, there have been some improvements to the air emissions mapping, and the flood risk mapping.

Changes in flood risk between WIMD 2008 and WIMD 2011 are largely due to the improved flood risk mapping. Although over 4,500 properties in Wales have benefited from reduced flood risk due to capital improvement and maintenance schemes since April 2005, half of these were in the last three years, and so not all of them will be included in the 2009 flood risk data used in WIMD 2011 and WIMD 2011: Child Index.

Publication	WIMD 2011 and WIMD 2011: Child Index
Indicator	Estimated LSOA Air Quality
Time Period	2008
Source	Environment Agency
Type	Score
Denominator	N/A (as it's a score)
Data availability	(StatsWales table)
Additional Notes	A 1km x 1km vector (polygon) grid was generated to cover Wales, corresponding to concentration GRIDs supplied by Netcen. Concentration values were then extracted for each LSOA for the following pollutants and statistics: Benzene annual mean Carbon Monoxide annual mean Ozone maximum daily 8 hour mean Nitrogen Dioxide annual mean Particulates annual mean Sulphur Dioxide annual mean 99.9th percentile of Sulphur Dioxide 15-minute means For land-based cells without data, concentration values were inferred using a simple average of surrounding cells. Figures for LSOAs were calculated by examining the overlap of the 1 km grid data with each LSOA and averaging the results of each grid intersecting the LSOA, weighted by the number of address points in the intersection. Each pollutant value was then adjusted to an equivalent scale using a factor based on the objective, standard or risk factor for that pollutant and statistic. The result is a population average for each LSOA. The number of households was provided by Royal Mail address point files. Particular exclusions made were: points outside Wales, PO Boxes, and Addresses containing either Organisation Name or Department Name. Although, inclusions made were if the organisation ends in: FARM, indicates HOSPITAL, indicates HOME, or indicates MOD Barracks. The method of combining data was developed to take into account air quality standards for each substance, which are based on the best medical and scientific understanding of their effects on health and/or the environment. The method also ensures that areas which have high prevalence of certain pollutants, but not others, are ranked as highly deprived; Low levels of one pollutant will not cancel out the effect of a high level of another pollutant. Poor air quality suggests proximity to certain activities such as traffic, domestic combustion and industrial sites – activities that could have a negative impact on quality of life, the local environment and health.

Publication	WIMD 2011 and WIMD 2011: Child Index
Indicator	Estimated LSOA Air Emissions
Time Period	2008
Source	Environment Agency
Type	Score
Denominator	N/A (as it's a score)
Data availability	(StatsWales table)
Additional Notes	A 1km x 1km vector (polygon) grid was generated to cover Wales, corresponding to concentration GRIDs supplied by Netcen. Concentration values were then extracted for each LSOA for the following pollutants: Arsenic Benzo[a]pyrene Butadiene Cadmium Chromium Dioxins Mercury Ammonia Nickel Nitrogen Oxides Lead Vanadium Volatile Organic Chlorides Figures for LSOAs were calculated by examining the overlap of the 0.5 km grid data with each LSOA and averaging the results of each grid intersecting the LSOA, weighted by the number of address points in the intersection. Each pollutant value was then adjusted to an equivalent scale using a factor based on objective, standard or risk factor for that pollutant. The result is a population average for each LSOA. The number of households was provided by Royal Mail address point files. Particular exclusions made were: points outside Wales, PO Boxes, and Addresses containing either Organisation Name or Department Name. Although, inclusions made were if the organisation ends in: FARM, indicates HOSPITAL, indicates HOME, or indicates MOD Barracks. The method of combining data was developed to take into account air quality standards for each substance, which are based on the best medical and scientific understanding of their effects on health and/or the environment. The method also ensures that areas which have high prevalence of certain pollutants, but not others, are ranked as highly deprived; Low levels of one pollutant will not cancel out the effect of a high level of another pollutant. While it is accepted that air quality is the preferred measure of risks from air pollution, air emissions data provides a good set of complimentary data covering pollutants not included in the Air Quality indicator. Emissions data are good indicators of proximity to polluting activities.

Publication	WIMD 2011 and WIMD 2011: Child Index
Indicator	Estimated LSOA Flood Risk
Time Period	2009
Source	Environment Agency
Type	Score
Denominator	N/A (as it's a score)
Data availability	(StatsWales table)
Additional Notes	Flood risk measures the proportion of people living in an area with a significant, moderate or low risk of flooding. Risk is based on frequency rather than level of damage of flooding. To ensure the areas at risk of more severe flooding rank as more deprived than areas at risk of less severe flooding, the following weighting was given: Number of households in an area at significant risk is multiplied by 0.04; number of households in an area at moderate risk is multiplied by 0.01; and number of households in an area at low risk is multiplied by 0.0025. Each of these numbers is calculated for each LSOA and then added together to give total normalised number of households at a risk of flooding per LSOA. This number is then divided by the total number of households in the LSOA to give the proportion of households at risk of flooding. The number of households was provided by Royal Mail address point files. Particular exclusions made were: points outside Wales, PO Boxes, and Addresses containing either Organisation Name or Department Name. Although, inclusions made were if the organisation ends in: FARM, indicates HOSPITAL, indicates HOME, or indicates MOD Barracks.

Publication	WIMD 2011 and WIMD 2011: Child Index
Indicator	Proximity to waste disposal and industrial sites
Time Period	2010
Source	Environment Agency
Type	Score
Denominator	N/A (as it's a score)
Data availability	StatsWales
Additional Notes	LSOAs are ranked based on the proportion of the population living within a 1km zone of each Pollution Prevention Control (PPC) site and active landfill site. Each site is assigned a band score which relates to the potential and the actual deprivation the site could cause to the environment and the people living within its vicinity. Operation Risk Appraisal (OPRA) scores were used to inform scores for sites which already have them. Sites which do not have OPRA scores were assigned scores by the Environment Agency. All mobile sites were excluded as their location may vary from day to day.

The following table shows how OPRA scores relate to site band scores:

OPRA Score	Site band score
>=500	5
200 - 499	4
100 - 199	3
75 – 99	2
50 – 74	1
< 50	0

The below matrix was used to create LSOA proximity scores:

LSOA distance from site (km)	Band score created from OPRA score (Industrial and waste management sites)					
	5	4	3	2	1	0
<=0.5	5	4	3	2	1	0
>0.5 to 1	4	3	2	1	0	0
>1 to 1.5	3	2	1	0	0	0
>1.5 to 2	2	1	0	0	0	0
>2 to 3	1	0	0	0	0	0
>3	0	0	0	0	0	0

There are instances where the 1km site zones overlap with each other. In these instances the proximity scores are added together. This means that the more sites an LSOA is close to, the more deprivation they will suffer.

To create the overall indicator score, an LSOA ranked percentile is needed. This is done by multiplying the number of address points within each site zone (regardless of LSOA boundary), by the sum of the proximity scores for these zones. All of these values are then summed to LSOA boundaries and divided by the number of address points within the LSOA to create the proximity to industrial or waste disposal sites indicator score for each LSOA.

The number of households was provided by Royal Mail address point files. Particular exclusions made were: points outside Wales, PO Boxes, and Addresses containing either Organisation Name or Department Name. Although, inclusions made were if the organisation ends in: FARM, indicates HOSPITAL, indicates HOME, or indicates MOD Barracks.

Appendix A Allocation of Data to Lower Layer Super Output Areas

Background

LSOAs are still a relatively new geography and as such data are generally not allocated to an LSOA as part of the collection process. Therefore for WIMD 2011, a method of allocating other geographical level data to LSOAs had to be devised.

Different data sets vary in their level of geographic coding and so different approaches are required. Data sets fall into one of the following categories, shown in order of preference in terms of data quality:

1. data are geocoded and can be allocated to LSOAs exactly using a Graphic Information System (GIS);
2. data contain the full postal address and can be allocated exactly to LSOAs using ONS lookup tables;
3. data are coded with some other small area geography and these can be allocated to LSOAs in some way, although the matching will not be exact.

The following sections explain the approach taken for WIMD 2008 in each case.

1 Geocoded data

Geocoding is still relatively rare particularly for national-level data sets. In the case of WIMD 2011 the only geocoded information available was the service point location information (e.g. schools, post offices etc.) used in the Geographic Access to Services
Domain.

2 Address Matching

This is the most accurate method for allocating data that has not been geocoded. The ONS have created a lookup to allocate each individual address to a LSOA which forms part of the toolkit for data providers of Neighbourhood Statistics (a package called MatchCode).

There are two main issues with this approach. The first is to do with address quality; addresses are often incomplete or partially incorrect (e.g. missing house number or incorrect postcode). There may then be a residual of records for which another approach is required (e.g. allocate on the basis of postcode). The second problem has to do with data confidentiality and legal restrictions which prevent the supply of full address information. The solution to the second problem is for the data provider to address match data and provide LSOA level counts and this was the case with benefits data that had already been coded for use in Neighbourhood Statistics.

3 Allocating small area data to LSOAs

The most common small area identifier is the postcode, and this was the building block used for many indicators in WIMD 2008.
Postcode level data

LSOA boundaries do not fit exactly with postcode boundaries. However, it was initially assumed that the fit would be good, that the occurrence of postcodes being split by LSOA boundaries would be minimal and that splits would generally result in the majority of a postcode clearly within a single LSOA. A detailed investigation found that there was a high proportion of split postcodes and that the

occurrence of postcodes effectively split down the middle was not negligible particularly in urban areas where LSOA boundaries can go down the middle of streets.

The standard approach with postcode level data is to allocate data on a best fit basis using the Postcode Address File (PAF), where postcodes are allocated to the area in which most of the population lies. As noted above, the match of postcodes to LSOAs in Wales is not a close fit and this causes concern with something like WIMD as rates are calculated using denominators from a different source and areas are then ranked. An alternative method of apportioning postcode data to each of the LSOAs that lie within the area was developed.

Apportionment of split Postcodes

The only information generally available with post code level data is the count (e.g. the number of deaths) and the rate that is derived from this. The basic principle used was that the rate for the given indicator should apply equally across the whole postcode and that this should be preserved under any allocation methodology. This can be done by weighting the postcode rates according to the proportion of the postcode population sitting within the LSOAs, as shown below:

$$LSOA\,rate = \frac{\sum \left(PC_{in}\ popn \times PC\ rate \right)}{LSOA\ popn}$$

Where $PC_{in}popn$ = population of the postcode within the LSOA

PC rate = postcode rate

LSOA rate = LSOA rate

LSOA popn = LSOA population

When postcodes are not split this essentially gives the rate a weighting of 1 as the postcode population figures cancel:

$$LSOA\,rate = \frac{\sum \left(PC\ popn \times PC\ num / PC\ popn \right)}{LSOA\ popn}$$

$$= \frac{\sum PC\ num}{LSOA\ popn}$$

$$= \frac{LSOA\,num}{LSOA\ popn}$$

where

PCpopn = total population of the postcode

PCnum = number of individuals in the postcode

LSOAnum = the total number of individuals in the LSOA.

When postcodes are split, it has been assumed that

$$\frac{PC_{in}\ popn}{PC\ popn} = \frac{Address\ Points_{in}\ num}{Address\ Points_{tot}\ num} = PercentagePCinLSOA$$

where

$AddressPoints_{in}num$ = number of postal address points in the postcode that are within the LSOA

$AddressPoints_{tot}num$ = total number of postal address points in the postcode

$PercentagePCinLSOA$ = estimated percentage of the postcode population within the LSOA

This says that the proportion of address points is equivalent to the proportion of the population, and hence assumes that population is distributed in the same way within each sub-area within the postcode. While it will not always be the case that the number of persons per household is the same in all parts of a postcode it will not generally be dramatically different. However, there is no realistic way to account for this (2011 Census data are too far out-of-date to be representative and are not available at a postcode level), and the WIMD Methodology Working Group agreed that the approach would provide more robust results than simple best fit.

Under this assumption, the LSOA rate is calculated as follows:

$$\begin{aligned} LSOA\,rate &= \frac{\sum\left(PC_{in}\ popn \times PC\ rate\right)}{LSOA\ popn} \\ &= \frac{\sum\left(\frac{PC_{in}\ popn \times PC\ num}{PC\ popn}\right)}{LSOA\ popn} \\ &= \frac{\sum\left(PercentagePCinLSOA \times PC\ num\right)}{LSOA\ popn} \end{aligned}$$

Which is equivalent to

$$LSOAnum = \sum\left(PercentagePCinLSOA \times PCnum\right)$$

Hence, approximate counts for indicators can be constructed at LSOA level by taking the relevant percentage of the counts for each postcode that falls within it. While this is the most straightforward way to estimate an LSOA figure, the above analysis shows that it is actually equivalent to giving every individual in split postcodes the rate for that postcode.

To facilitate this approach the Welsh Assembly Government Cartographics Unit produced a look up table for 2007 postcodes, where it was possible to calculate apportionment rates as above for each postcode. 2006 and 2005 look up tables were also supplied by the Welsh Assembly Government Cartographics Unit.

Best Fit to Postcode Boundaries

While the apportionment method has been used for most WIMD postcode level data, there were a small number of instances where out-of-date postcodes were contained in data records and the only option was to use the best fit method. A best fit method could be used as the ONS All Fields Postcode Directory (AFPD) holds historical postcodes.

Appendix B Population Denominators

Background

For the Welsh Index of Multiple Deprivation (WIMD) 2011, population estimates were needed for each of the 1,896 Lower layer Super Output Areas (LSOAs) to provide population denominators for many of the individual indicators. Estimates for the 'at risk' population were constructed using a combination of LSOA Small Area Population Estimates (SAPE) and LSOA level 2001 Census populations, both published by the Office for National Statistics (ONS).

The population estimates are on a usual residence basis and as such include household population and communal establishment population (e.g. students, persons in care establishments and children in local authority homes) but the prison population has been excluded. Prisoners are a special case in that they are not at risk for some forms of deprivation but are in any case isolated from the rest of society and as such do not contribute to the overall deprivation level.

SAPE LSOA figures have been used in WIMD 2008 and 2011.

Lower layer Super Output Areas

Lower layer SOAs (used for the WIMD 2011) contain a minimum population of 1,000 and a mean population of 1,500. These typically contain between 4 and 6 Census Output Areas, and are constrained by the boundaries of the Standard Table (ST) wards used for the 2001 Census. In Wales there are 1,896 LSOAs.

Other denominators

For some of the indicators, denominators were taken from the same datasets as the numerators. Indicators for which this applied were denominators for:

- The Key Stage 2, 3 and 4 indicators in the Education Domain.
- Absenteeism indicators for the Education Domain.
- No Qualifications indicator for the Education Domain.
- Limiting Long Term Illness from the Health Domain.
- Singleton Low Birth Weight from the Health Domain
- All indicators for the Housing Domain.
- All denominators for the Geographical Access to Services Domain were the numbers of address points within each LSOA.

Appendix C The Factor Analysis Technique

Factor Analysis Overview

Factor analysis is a method for assessing the extent to which a set of indicators may be measuring the same underlying construct or factor. The premise behind a one-common-factor model is that the underlying factor is imperfectly measured by each of the indicators in the dataset but that indicators that are most highly correlated with the underlying factor will also be highly correlated with each other. By analysing the correlation between indicators it is therefore possible to make inferences about the common factor and as a result estimate a 'factor score' for each LSOA. This score is derived from a set of weights for each of the indicators in the data set that is generated by the process of factor analysis. This factor score can then be used as the domain index.

Factor analysis has only been applied to four domains: Health, Education, Geographical Access to Services and Community Safety. Factor analysis is used in these domains because they contain indicators that measure, on potentially different metrics and with different levels of accuracy, a number of forms of that deprivation and therefore cannot otherwise easily be combined. The main reasons why Factor Analysis has been used are:

- Because the indicators are on different metrics and have different levels of accuracy, and so cannot simply be summed
- To ascertain the factor that underlies the indicators within the Domain
- To help take into account the problem of 'double counting' within a Domain

In the Employment and Income domains we can identify individuals who are or are not deprived in terms of the domain definition. The number of deprived people can then simply be summed and divided by a suitable denominator to create an area rate. This is not possible in the other six domains. These deprivations tend to present themselves in different ways at different times. Thus, for example, an individual is 'health deprived' if they die prematurely or are long-term sick. While the long-term sick may be more likely to die prematurely than others, these events do not occur to the same people at the same time. Typically such domains include data on people at different ages and stages e.g. in the education domain, lack of qualifications in the adult population as well as poor results at school level. Instead we hypothesise that there is an underlying factor at the local area level (e.g. health deprivation) that makes these different states likely to exist together in the same area. This underlying factor cannot be measured directly but can be identified through its effects on specific individual measures (e.g. premature death, long-term sickness, low birth-weight children etc.). We have therefore collected a number of indicators that measure, with different levels of accuracy, the effects of this underlying factor. By looking at the relationship between all these indicators the underlying factor can be identified and quantified.

Factor analysis also takes some account of the problem of 'double-counting' within domains. The Health, Education, Access to Services and Community Safety domains potentially contain indicators that overlap with each other. For example, in the Health domain, it is possible for an individual to have had cancer and also potentially to be included in the limiting long-term illness indicator. Combining data using other methods such as 'z scores' more directly double-weights these cases by taking them all into account. Factor analysis, however, takes some account of this overlap in that an indicator may have a lower weight if the contribution it makes has already been taken into account.

The choice of maximum likelihood estimation method

WIMD 2011 follows the WIMD 2005 and WIMD 2008 methodology and that applied by Oxford University for WIMD 2000 as well as the Indexes for the other three UK countries.

In Principal Components Analysis all variance in an indicator is analysed including measurement error (*error variance*) and the indicators' imperfect measurement of the underlying construct or constructs (*specific variance)*. This is because it does not attempt to separate *common variance* (i.e. variance shared between three or more indicators) from *unique variance* (i.e. specific variance and error variance). It assumes that an indicator is perfectly reliable and measured without error. It was therefore not appropriate to use the Principal Components method. The appropriate technique, where it is suspected that indicators are not perfectly reliable or measured without error, is *common factor analysis* of which Maximum Likelihood Factor (ML) analysis is a type.

Principal Factoring (PF) has, in the past, been the favoured method of common factor analysis but this was probably because of its relative computational simplicity. With the advent of high-powered computers more sophisticated methods, such as ML factor analysis, are now easily accomplished. PF has a number of disadvantages in comparison to ML factor analysis. The PF solution depends on the scale of measurement of the input indicators (i.e. depends on whether or not they have been standardised), which means that there is not one but an infinity of PF solutions, the choice among, which is arbitrary. The factor model itself is intrinsically scale free, and thus any procedures for its estimation should be scale invariant. ML is scale invariant. ML also treats the correlation matrix as a sample correlation matrix and attempts to explain variance in the *population* correlation matrix. This treatment of the data as a sampled dataset is consistent with the proposal, made throughout this project, that even 'census' indicators should be seen as a sample from a super-population.

Communality

This is the proportion of a variable's variance explained by a factor structure. A variable's commonality must be estimated prior to performing a factor analysis. A communality does not have to be estimated prior to performing a principal component analysis. Communality estimates are estimates of the proportion of common variance in a variable. *Prior communality estimates* are those which are estimated prior to the factor analysis. Common methods of prior communality estimation are to use (1) an independent reliability estimate, (2) the squared multiple correlation between each variable and the other variables, (3) the highest off-diagonal correlation for each variable, or (4) iterate by performing a sequence of factor analyses using the final communality estimates from one analysis as prior communality estimates for the next analysis. *Final communality estimates* are the sum of squared loadings for a variable in an orthogonal factor matrix.
The default setting for communality prior estimates, Square Multiple Correlation, was used for WIMD 2011 calculations.

Calculation Process

The indicators were first transformed to the standard normal distribution. The transformed indicators were then entered into a 'one common factor Maximum Likelihood factor analysis'. Fuller's regression method was used to derive factor scores from the resulting solution. The process was undertaken in SAS and the following details the settings used.

- The normally transformed values for each of the domain indicators were entered as the analysis variables;
- Maximum likelihood factor analysis was chosen as the factoring method, for the reasons described above;
- The smallest eigenvalue was set to 1 because this is a commonly used indicator showing that sufficient factors have been extracted to reasonably explain the 'common variance' between the indicators.

- For prior communality estimates the method chosen was Squared Multiple Correlation with all other columns, as described above;
- For the rotation method, no rotation was selected as we are only looking for a single factor solution and rotation only applies if there two or more factors.

Appendix D Exponential Transformation of the Domain Indexes

The precise transformation involved is as follows. For any LSOA, denote its rank on the domain, scaled to the range [0,1], by R (with R=1/1896 for the least deprived, R=1896/1896=1 for the most deprived).

The transformed domain, (X) equals:

-23*log{1 - R*[1 - exp(-100/23)]}

where log denotes natural logarithm and exp the exponential or antilog transformation, and * denotes multiplication. This formula is straightforwardly calculated and is in fact simpler than the commonly-used transformation to a normal curve which necessitates the use of a look-up table. The resulting distribution is illustrated below in a histogram.

Each transformed domain has a range of 0 to 100, with a score of 100 for the most deprived LSOA. Ten percent of LSOAs have a score higher than 50. When transformed scores from different domains are combined by averaging them, the skewness of the distribution reduces the extent to which deprivation on one domain can be cancelled by lack of deprivation on another. For example, if the transformed scores on two domains are simply averaged, with equal weights, a (hypothetical) LSOA that scored 100 on one domain and 0 on the other would have a combined score of 50 and would thus be ranked at the 90th percentile. (Averaging the untransformed ranks, or after transformation to a normal distribution, would result in such a LSOA being ranked instead at the 50th percentile: the high deprivation in one domain would have been fully cancelled by the low deprivation in the other.) Thus the extent to which deprivation in some domains can be cancelled by lack of deprivation in others is, by design, reduced.

Histogram of a transformed domain