**TAGRA Morbidity & Life Circumstances (MLC) Subgroup**

**Review of the MLC adjustment for the Acute care programme in the NRAC formula**

**FINAL REPORT**

**NHSScotland Information Services Division (ISD)**

**Scottish Government (Health Analytical Services Division)**

*August 2016*

## ACKNOWLEDGEMENTS

The subgroup wishes to thank the following people for their assistance with the work of this review.

David Baird and Julie Peacock, NHS ISD, gave a presentation on the Patient-Level Information Costing System (PLICS).

Colin Fischbacher and David Brewster, NHS ISD, gave clinical advice to analysts on particular questions.

Bruce Guthrie, University of Dundee, presented his work on multimorbidity to the subgroup.

Ali Jarvis, NHS Health Scotland, and Andrew Waugh, Scottish Government, came to a meeting to provide their expertise on health inequalities.

Alastair McAlpline, Scottish Government, gave advice to analysts on the practicalities of using the Scottish Index of Multiple Deprivation (SIMD).

Barry Melia, NHS ISD, gave advice to analysts on the use of prescribing data in candidate indicator variables.

Debbie Sigerson and Emma Doyle, NHS Health Scotland, came to meetings to assist in the HIIA process.

Matt Sutton, University of Manchester, participated in discussions around the geography decision and advised on unmet need methodology.

Bruce Whyte, Glasgow Centre for Population Health, and Gerry McCartney, NHS Health Scotland, advised on population inequality data.

## EXECUTIVE SUMMARY

### Background, MLC overview and remit of review

The NHSScotland Resource Allocation Committee (NRAC) formula is used to inform the allocation of funding to the 14 territorial NHS Boards for Hospital and Community services and GP Prescribing, accounting for around 70% of NHSScotland’s budget. It was developed in 2007 from the previous Arbuthnott formula [1].

The Technical Advisory Group on Resource Allocation (TAGRA) was set up to oversee ongoing maintenance and development of the formula. The work of the Acute Morbidity and Life Circumstances (MLC) subgroup presented in this report is part of a rolling programme of review.

The MLC adjustment takes account of additional needs over and above those based on age and sex. Areas where residents have greater levels of ill health, or are subject to life circumstances that result in ill health, will impose higher costs on NHS Boards in meeting the increased volume of healthcare activity needed by those populations. The adjustment is calculated through a linear predictive model, based on the factors that best explain the variation in age-sex-adjusted costs of healthcare between small areas for each care programme. (See section 1.2 of the main report for further details on the MLC calculation.)

For the year 2016/17, the Acute care programme accounted for around £4.5billion, 52.4% of the total £8.5billion allocated via the NRAC formula. The MLC subgroup were asked to recommend to TAGRA changes to the Acute MLC adjustment, to improve the ability of the formula to allocate funds between the territorial NHS Boards on a fair and equitable basis, having regard to TAGRA’s core criteria (see section 1.4). In particular, the group was asked to consider:

* Costing method – to assess the effects of a change in costing method (should that be agreed by TAGRA) on the adjustment factors;
* Time – the degree to which observations should be aggregated over multiple years, in order to provide stable adjustment factors;
* Geography – the appropriate geographical level (data zones or intermediate zones) at which to carry out the adjustment;
* Structure – the form, the age grouping and clinical scope, specified for the regression model that will create the adjustment factors;
* Indicators – the most appropriate indicators to use within any adjustment, including but not limited to indicators of need, indicators of supply (for control purposes), rurality, and urbanity.

### Costing method (chapter 3)

The calculation of the MLC adjustment is based on (implicit) estimates of the cost of individual episodes of care. In the NRAC costing methodology, for inpatients, the episode costs include both a *fixed* and a *variable* element. Fixed costs represent costs such as overheads, theatres, and medical staff which are applied to all episodes equally – a “cost of admission”. Variable costs represent direct costs such as nursing staff and pharmacy which are applied on the basis of length of stay of the patient – a “cost per day”. This is known as the “case-mix adjustment”: the cost for each episode depends (partially) on the length of stay.

The fixed and variable costs per episode of care are calculated as follows: First a ratio is applied to total specialty costs, to divide that total into a fixed and a variable component. The fixed component is divided by the number of episodes of care in the specialty to estimate a standard fixed cost for each episode of care in that specialty. The variable component is divided by the number of days of stay in that specialty to estimate a standard cost of care per day. These two elements are combined to estimate a cost for each individual episode of care: the specialty cost per day multiplied by the length of stay for that individual episode is added to the fixed cost per episode for that specialty.

The ratios used to divide costs into fixed and variable for each specialty were derived in the 2007 NRAC review from a regression analysis of the relationship between average cost per episode and average length of stay (by hospitals) within each of five specialties.

It was decided to retain the NRAC costing method, but to replace the regression-derived ratios for dividing specialty costs into fixed and variable costs with ratios derived from the new Patient-Level Information Costing System (PLICS) episode-level costs, produced by ISD’s Integrated Resource Framework (IRF) team. The new ‘fixed’ and ‘variable’ proportions for each specialty are derived by summing the relevant PLICS direct cost components for each episode of care.

### Model specification (chapter 4)

Functional form: the subgroup decided to retain the linear modelling approach currently used, mainly on grounds of transparency and practicality.

Time span: the subgroup decided that using 3 years of data for the dependent variable (the ratio of the actual cost in a small area to the expected cost based on age and sex) was preferable to using just a single year. Using 3 years will smooth out random year-to-year variations which would be likely to improve stability, while remaining sufficiently responsive to any incremental changes in the relationships over time.

Geography: the subgroup decided that the Acute MLC adjustment should be calculated at data zones rather than intermediate zones. Finer granularity should lead in principle to better equity and responsiveness, and in a large care programme such as Acute, there is arguably no reason to aggregate data to larger geographical units.

Diagnostic groups: the current Acute MLC adjustment analyses cost variation within seven diagnostic groups: Cancer, Heart, Respiratory, Digestive, Injury, Outpatients and Other. The subgroup preferred to retain diagnostic groups, rather than combining them together, on grounds of face validity, possible improvement to the predictions at local level, and relevance to a range of stakeholders.

Age split: the subgroup considered whether the Acute MLC adjustment could be improved by introducing an age split, as had been done for Mental Health and Learning Difficulties in the recent review of its MLC component [2]. However, analysis did not show an obvious point for the split, and advice from ISD clinical consultants was that there are no confident clinical grounds on which to introduce a split by age group. The subgroup decided that the indicator selection process would be done based on an all-ages model, and then subsequent analysis would be carried out using several trial age splits to check the performance of the model across age groups. This analysis showed that at the “all ages” level, there is very little difference between the predictive power of models with or without an age split. The recommendation is therefore to retain the Acute MLC all-ages model.

### Supply model (chapter 5)

The candidate supply variables that were previously explored in the 2007 NRAC review were re-examined in the current Acute MLC review. The analysis showed that the supply variables currently in use – IPACX and OPACX – were not outperformed by any other variables, so the subgroup was content to retain them. However, it was decided that since these two variables are highly correlated, it would be better to use just one of them: IPACX for inpatient diagnostic groups, since it relates to proximity of inpatient facilities; and OPACX for Acute Outpatients, since it is based on outpatient facilities.

### Needs index development (chapter 6)

A selection process was carried out using 58 candidate indicators (Table 7). The list of previously-tested indicators formed the starting point for this, with additional variables included based on the subgroup’s knowledge of new data sources and in consultation with experts in NHSScotland. The subgroup agreed to consider only indicators which had a theoretical link to Acute health need (for transparency and face validity reasons), that had a geographical granularity at least as low as intermediate zones, for which the available data was no more than 10 years old, and that were able to be updated going forward.

The Acute MLC review followed broadly the same indicator selection methodology as used in the original NRAC review. However, since the indicators of need were eventually to be combined into a single ‘needs index’ (as the sum of the Z-scores of the indicators) for use in the NRAC formula, it was agreed that the selection process should look for the best-performing *index* options, rather than constructing indices from the best-performing sets of *separate* variables (as had been done in 2007).

The analysis identified the ‘best’ single-variable index and the best two-, three- and four-variable index options for the subgroup to consider. Adjusted R2 was used as the primary metric to compare the index options’ performance: R2 is a common goodness-of-fit measure representing the proportion of the dependent variable’s variance that is explained by the model; “adjusted” R2 adjusts for differing model complexity i.e. different numbers of variables. Predictive power was also evaluated, by comparing the predictions of each fitted model with a 1-year cost ratio based on more recent data than that used to fit the model; the Residual Sum of Squares (RSS) was used as the metric for predictive power.

The possibility of different needs indices for different diagnostic groups was considered. However, there was no real evidence that different condition groups have different needs drivers (at least, from the set of candidate variables analysed in this review), nor that diagnostic group-specific indices would perform better over the long term and justify the extra complexity. The subgroup therefore decided that the Acute MLC adjustment should continue using a common index across all diagnostic groups.

The selection process led to a conclusion that the best option for the Acute MLC index remains the two-variable option selected in the 2007 NRAC review:

* All-cause Standardised Mortality Ratio (SMR) for ages 0-74. This is a measure of premature deaths.
* Limiting long-term illness (LLTI) rate. This is an age-sex standardised rate of self-reported limiting long-term illness from Scotland’s census.

### Additional variables (chapter 7)

A number of outliers were apparent in scatter plots, for which the cost ratios were substantially larger than would be expected from the linear trend; this was particularly the case for Outpatients. Investigation of the two largest cost ratio outliers for Outpatients revealed that the data zones both contained prisons. The subgroup considered whether the few data zones containing prisons should be removed from the regression analysis. However, because the funding of Acute healthcare for prisoners is covered by the NRAC formula, the subgroup felt this additional need should be captured in the MLC model. The recommendation is to include a ‘dummy’ variable (i.e. a binary indicator) to represent the presence of a prison within a data zone, so that the additional cost associated with prisoner Acute care could be modelled (simply, as a constant amount to be added to a data zone containing a prison) and included in the target shares. This should be done for Acute Outpatients alone since the dummy variable turned out not to be significant for other diagnostic groups.

Possible urban-rural effects were investigated; following the method of previous reviews, this was done by including urban-rural (UR) markers as extra binary variables in the regression of cost ratios upon the needs index and supply model, which would effectively adjust the healthcare need of each area within the same urban-rural category by a constant amount. The explanatory power was increased only very slightly when urban-rural markers (of any classification) were included. The predictive power also did not improve substantially when including the urban-rural markers. The inclusion of urban-rural markers did not produce significant differences in the coefficient of the proposed Acute needs index. The subgroup members were in agreement that the proposed Acute MLC model appears to perform similarly well across all urban-rural settings and no significant improvement could be anticipated from including urban-rural indicators in the model.

### Unmet need (chapter 8)

Shortfall methods have been used to check for evidence of unmet need in the form of lower-than-expected utilisation (based on the linear model) in the data zones with the highest needs index values, the most urban zones, the most rural zones, and the data zones with the highest proportion of ethnic minority populations.

The aim of shortfall methods is to help identify whether, in a given subset of data zones, the utilisation is lower than it should be due to unmet need. If these affected data zones influence the slope of the fitted linear model, they can be excluded from the regression in the eventual implementation of the MLC model. The resulting ‘adjusted’ slope is then used to predict need for *all* data zones. There is currently an unmet need adjustment for Heart based on excluding the 25% most deprived intermediate zones, based on the Scottish Index of Multiple Deprivation (SIMD) income domain, from the regression.

The present review found that there were no significant shortfalls relating to urban / rural setting or ethnicity. Shortfalls were, however, found for data zones with high values of the needs index, for Heart and Other diagnostic groups. The optimal cut-off point was 30% for Heart (i.e. the data zones with the highest 30% of needs index values should be removed from the regression), and 5% for Other.

### Health Inequalities Impact Assessment (Chapter 9)

The subgroup aimed to embed equity considerations throughout the Acute MLC review and adapted NHS Health Scotland’s HIIA materials [14] to devise a tailored process to support systematic consideration of equity throughout. This comprised three components: an equality advisor to join the subgroup and provide regular, informal assessment of proposals and analyses at and between meetings; a focussed HIIA workshop discussion at the midpoint; and a process of agreeing recommendations for equity-related variables following further analyses. The first component resulted in expanding the definition of equity within the review’s core principles, and the identification of a number of potential explanatory variables for the Acute MLC model – of which life expectancy was considered but rejected, and DNA rates reached the final candidate list. The second component was an HIIA workshop discussion facilitated half way through the review to which additional equality advisers were invited, similar to that used for a more formalised HIIA process although carried out an earlier stage. The workshop identified ethnicity and unpaid care as candidate variables and these were put forward for further analyses. The third component comprised a process of reaching consensus on recommendations from the HIIA process.

Consensus was reached as follows: DNA rate was retained as a candidate variable until a late stage in the process, but was rejected due to the potential for creating a perverse incentive for Boards as a supply variable (described in section 6.1.12); unpaid care was explored as a candidate variable, but was rejected due to the absence of a dose-response relationship between unpaid care provision and higher acute costs (section 6.1.7); ethnicity was the subject of a number of analyses (described in Chapter 6 and Chapter 8), but was eventually rejected as an explanatory variable as the complexity and depth of need could not be captured in a way that would remain stable over time and / or that would show a level of impact that would require adjustment to the formula. The subgroup agreed that unpaid care and ethnicity should be re-examined in later reviews if better ways of capturing need could be developed, and that there may be a need for other ways to think about health care provision for small groups, for example, through the High Resource Individuals initiative or through the Scottish Allocation Formula (SAF) for GP practices. While the HIIA process did not identify new indicators for inequality, the process modelled a commitment to analysing the impact of the formula on equity in acute healthcare. The benefits of considering equity throughout the process were that there were multiple opportunities to consider equality, inequality and human rights throughout the review and that the whole subgroup and review team were involved.

### Further recommendations (chapter 10)

Following the conclusion of the Acute MLC subgroup, in addition to the main recommendations below, the subgroup recommends that the following are considered:

* PLICS Acute costs should be introduced in future revisions of the NRAC formula.
* Given the research findings on importance of multimorbidity and mental health conditions in predicting need for acute services, better data are needed at data zone level to be able to assess these issues. In particular, primary care data from the Scottish Primary Care Information Resource (SPIRE) should be explored when it becomes available; prescribing data could also be further explored and linked to other data sets to provide information on additional co-morbidities. Improvement in the recording of co-morbidities in SMR01 should also be utilised.
* Better data and methods should be developed to capture need related to specific ethnic groups and those providing unpaid care.
* As the level of unmet need for cardiac services has not improved since the last revision of the formula, policy measures are needed in addition to the formula adjustment recommended here.

### Review outcomes

The main recommendations resulting from the Acute MLC review are listed below.

|  |
| --- |
| Recommendation 1: retain the NRAC costing method, but with the case-mix adjustment derived from summing PLICS components by specialty. |
| Recommendation 2: retain the linear functional form. |
| Recommendation 3: change the time span of the Acute MLC cost ratios from 1 year to 3 years. |
| Recommendation 4: change the geography basis of the Acute MLC adjustment from intermediate zones to data zones. |
| Recommendation 5: retain the Acute diagnostic groups. |
| Recommendation 6: continue to model all ages together for the Acute MLC adjustment. |
| Recommendation 7: continue to use IPACX and OPACX as the supply variables, along with the health board dummies, but use *only* IPACX for inpatient diagnostic groups and *only* OPACX for the Outpatients diagnostic group. |
| Recommendation 8: retain Limiting Long-Term Illness and All-cause Standardised Mortality Ratio ages 0-74 as the needs indicators. |
| Recommendation 9: introduce a prison indicator (a binary variable) for the Outpatients diagnostic group, to allow additional need associated with prisons to be allocated for. |
| Recommendation 10: as for the current formula, do not add urban-rural category indicators in the Acute MLC model. |
| Recommendation 11: implement two unmet need adjustments for   * the Heart diagnostic group by excluding the 30% of data zones with the highest needs index values from the regression, * the Other diagnostic group by excluding the 5% of data zones with the highest needs index values from the regression. |

The proposed changes to the Acute MLC adjustment are summarised as follows:

|  |  |  |
| --- | --- | --- |
|  | **Current Acute MLC adjustment** | **Proposed new Acute MLC adjustment** |
| Geography | Intermediate zones | Data zones |
| Time span | One year’s data is used to calculate the cost ratios | Three years’ data is used to calculate the cost ratios |
| Diagnostic groups | Heart, Cancer, Respiratory, Digestive, Injuries & Poisoning, Acute Outpatients and Acute Other | No change |
| Needs indicators | All-cause Standardised Mortality Ratio (SMR) for ages 0-74  Limiting long-term illness (LLTI) rate | Same needs indicators, plus an extra (binary) prison indicator for Acute Outpatients |
| Supply variables | IPACX and OPACX  Health Board dummies | IPACX for inpatient diagnostic groups, OPACX for Acute Outpatients  Health Board dummies |
| Unmet need | Adjustment for Heart based on excluding the 25% most deprived intermediate zones from the regression, as defined by SIMD income domain | Adjustment for Heart based on excluding the 30% of data zones with the highest needs index values and for Other based on excluding the 5% of data zones with the highest needs index values. |

## CHAPTER 1: BACKGROUND AND INTRODUCTION

### 1.1 Background

The NHSScotland Resource Allocation Committee (NRAC) formula is used to inform the allocation of funding to the 14 territorial NHS Boards for Hospital and Community services and GP Prescribing, accounting for around 70% of NHSScotland’s budget. The formula uses a weighted capitation approach to distribute funds equitably that takes into account need for healthcare related to demographics (age and sex) and additional need (morbidity and life circumstances), as well as taking into account unavoidable excess costs to supply services in remote areas.

The formula was developed in 2007; in their final report [1], NRAC recommended ongoing maintenance and development of the formula. The Technical Advisory Group on Resource Allocation (TAGRA) was set up for this purpose. The work of TAGRA’s Morbidity and Life Circumstances (MLC) subgroup is part of a rolling programme of review.

Following the completion of the review of the MLC component for the Mental Health and Learning Difficulties care programme in December 2012 [2], it was agreed that the Acute care programme’s MLC adjustment would be reviewed next and that the detailed analysis required would be undertaken by a technical subgroup.

The Acute care programme accounted for around £4.5bn, 52.4% of the total £8.5 billion allocated via the NRAC formula for the year 2016/17. It is by far the largest of the care programmes (for comparison, the next largest care programme is Mental Health and Learning Difficulties, accounting for 10.8% of the total).

The Acute care programme is broken down into seven diagnostic groups for the MLC adjustment, based on ICD-10 code groups. These, along with activity and cost information for each, are shown in Table 1.

Table 1: Acute diagnostic groups

|  |  |  |
| --- | --- | --- |
| **Diagnostic group** | **2013/14 Number of episodes** | **2013/14 Actual spend** |
| Cancer | 199,940 | £414m |
| Heart | 157,121 | £404m |
| Digestive | 189,768 | £343m |
| Injury | 124,769 | £407m |
| Respiratory | 135,354 | £291m |
| Outpatients | 1,466,760 | £769m |
| Other | 727,869 | £1,368m |
| **Total** | **3,001,581** | **£3,997m** |

### 1.2 NRAC’s MLC component

The MLC adjustment takes account of additional needs over and above those based on age and sex. Areas where residents have greater levels of ill health, or are subject to life circumstances that result in ill health, will impose higher costs on health boards in meeting the increased volume of healthcare activity needed by those populations.

The adjustment is calculated through a linear predictive model, based on the factors that best explain the variation in healthcare utilisation between small areas for each care programme. The utilisation of healthcare is represented by the ratio of actual costs (taking into account activity type and length of stay in that specific neighbourhood) to the expected costs (based on the neighbourhood’s population structure and national age/sex average cost per head).

A single explanatory variable is constructed from the indicators of need for each care programme, using the sum of the Z-scores of the individual variables. This combined variable is referred to as the “needs index”. Health board ‘dummy’ variables and supply variables are included in the regressions, but not in the predictions, to control for effects that are largely due to variations in supply.

The original work undertaken for the NRAC Committee, which resulted in the current MLC adjustment, is outlined in Chapter 5 of the NRAC report, and set out in detail in Technical Report D [3] and the Addendum to Technical Report D [4].

### 1.3 Current Acute MLC adjustment

The basis of the current MLC adjustment for the Acute care programme is:

* Geography: the adjustment is applied at intermediate zone level (there are 1,235 intermediate zones within Scotland), as for all other care programmes.
* Time span: one year’s data is used to calculate the cost ratios.
* Needs indicators: the indictors of need currently used for the Acute care programme are:
  + All-cause Standardised Mortality Ratio (SMR) for ages 0-74. This is a measure of premature deaths.
  + Limiting long-term illness (LLTI) rate. This is an age-sex standardised rate of self-reported limiting long-term illness from Scotland’s census.
* Diagnostic groups: the linear model is calibrated to cost ratios for 7 separate diagnostic groups: Heart, Cancer, Respiratory, Digestive, Injuries & Poisoning, Acute Outpatients and Acute Other. The predictions within the diagnostic groups are then aggregated together.
* Unmet need: there is an unmet need adjustment for Heart based on excluding the 25% most deprived intermediate zones from the regression, as defined by the Scottish Index of Multiple Deprivation (SIMD) income domain.

### 1.4 Remit and terms of reference for the Acute MLC review

The following remit and terms of reference for the technical subgroup were agreed:

To recommend to TAGRA changes to the Acute MLC adjustment, to improve the ability of the formula to allocate funds between the territorial NHS Boards on a fair and equitable basis, having regard to TAGRA’s core criteria (see below).

In particular, the group was asked to consider:

* Costing method – to assess the effects of a change in costing method (should that be agreed by TAGRA) on the adjustment factors;
* Time – the degree to which observations should be aggregated over multiple years, in order to provide stable adjustment factors;
* Geography – the appropriate geographical level (data zones or intermediate zones) at which to carry out the adjustment;
* Structure – the form, the age grouping and clinical scope, specified for the regression model that will create the adjustment factors;
* Indicators – the most appropriate indicators to use within any adjustment, including but not limited to indicators of need, indicators of supply (for control purposes), rurality, and urbanity.

TAGRA agreed that the work of the Acute MLC subgroup should take around 2 years (2014-2015). This was later extended by 8 months due to data availability issues; see chapter 2 for details.

The subgroup was appointed to have a wide membership to reflect the diversity of stakeholder interests in the Acute Morbidity and Life Circumstances (MLC) adjustment. A full list of members can be found in Annex A. The subgroup agreed to report on its progress at each meeting of TAGRA and seek TAGRA’s input on all aspects of the analysis.

The analysis for the review was primarily carried out by ISD members, with support from Scottish Government ASD members. These two groups comprised an Analytical Support Team (AST) which met regularly outwith the subgroup meetings. Quality assurance was provided throughout the review’s analytical work by the use of self- and peer-checking procedures, scrutiny of results by the subgroup, and full investigation of any unexpected or unusual findings.

A Health Inequalities Impact Assessment (HIIA) was integrated into the work of TAGRA from the outset, as described in Chapter 9.

The subgroup met 16 times over the period February 2014 to July 2016. All meeting papers and minutes were publicly available at <http://www.tagra.scot.nhs.uk/subgroups/morbidity-and-life-circumstances/>. Papers referred to throughout this report can be found at this URL.

Throughout the work, adjusted R2 was used as the primary metric to compare different models’ performance: R2 is a common goodness-of-fit measure representing the proportion of the dependent variable’s variance that is explained by the model; “adjusted” R2 adjusts for differing model complexity i.e. different numbers of variables. To evaluate the models in predictive mode, predicted cost ratios from each model – generated using the standard NRAC method – were compared with a 1-year cost ratio based on more recent data than that used to fit the model. Predictive power was then evaluated using the residual sum of squares (RSS): the sum of the squared differences between the predictions and the observations.

The subgroup adopted the same set of core criteria as used by TAGRA, and used them to frame discussion and aid decision-making, when deciding between analytical options. However, as a result of the Health Inequalities Impact Assessment (HIIA) process (see Chapter 9), the equity criterion was reworded to expand the definition of equity to specifically mention variation in need across population groups.

The final core criteria used by the subgroup and then adopted by TAGRA are as follows:

|  |  |
| --- | --- |
| **Equity** | The primary consideration should be to achieve the greatest possible accuracy in capturing the cost implications of variations in need between population groups and across the country, in order to develop a formula that delivers the greatest possible equity of access to health services. |
| **Practicality** | Use should be made of good-quality, routinely-collected data, in order to produce an administratively feasible formula that can be readily updated. |
| **Transparency** | The rationale informing the formula’s methodology should be explicable and any judgements should be made explicit, although this should not lead to over-simplification of details which might add precision to the methods. |
| **Objectivity** | The formula should as far as possible be evidence-based, using as necessary the full range of available robust data. |
| **Avoiding perverse incentives** | The formula should guard against perverse incentives and any negative consequences which might threaten the integrity of the data. |
| **Relevance** | There is a need to avoid the dangers of extrapolation and to make explicit where hard information is being used about one aspect of a service to make some assumption about an area where information is less good or absent. |
| **Stability** | There should be a reasonable degree of year-to-year stability in the data sources feeding in to the formula. |
| **Responsiveness** | The formula should result in shifts in the allocation of resources in response to changes in the need for healthcare services. |
| **Face validity** | The outcome of any changes to the formula should be subjected to a 'common-sense' check. |

The structure of this report is as follows. Chapter 2 discusses the data used, including the decision to carry out the bulk of the review at the 2011 geographies and its implications. Chapter 3 describes the review of the costing methodology. Chapter 4 discusses various aspects of the model specification that were reviewed: functional form, time span, geography, diagnostic groups, and the possibility of an age split. Chapter 5 provides information on the supply models considered and the selection process for the supply model. Chapter 6 presents the indicator selection process: candidate variables, methodology, and results. Chapter 7 mentions additional binary variables that were explored for inclusion in the MLC prediction model; namely, a prison indicator (which was adopted for Acute Outpatients), and urban-rural category indicators (which were tested but not adopted). Chapter 8 describes the process of checking for evidence of unmet need using shortfall methods. Chapter 9 describes the Health Inequalities Impact Assessment process as implemented in the Acute MLC review. Chapter 10 highlights various limitations of the data available for use, and outlines further recommendations for the future.

## CHAPTER 2: DATA USED

### 2.1 Issues around redrawn geographies

On 6th November 2014, the Scottish Government released the new ‘2011’ data zone and intermediate zone geographies following the release of the 2011 Census results. This represented an extensive change from the ‘2001’ geographies, with boundary changes for around 40% of the data zones, and a similar proportion of intermediate zones. At this time, the subgroup had to decide whether to carry out the remainder of the Acute MLC review analysis – including the important indicator selection work – at the new 2011 geographies, or at the old 2001 geographies currently used in the formula.

The extensive changes reflected extensive population migration over time. Originally, data zones were intended to have populations between 500 and 1,000 residents and intermediate zones between 2,500 and 6,000 residents. In 2005, 683 (10%) of data zones and 57 (5%) of intermediate zones had populations outside of these parameters. In 2013, this had increased to 1,140 (18%) of data zones and 130 (11%) of intermediate zones. This movement perhaps implies that the original socioeconomic homogeneity within each zone had also been altered.

The subgroup considered the implications of this for the modelling of healthcare utilisation in small areas. Paper TAMLC22 examined the options around geographies. It was generally felt that better and more “future-proof” results could be expected from using the new geographies. However, this had to be balanced against data availability issues at the new geographies.

Most significantly, the National Records of Scotland (NRS) mid-year population estimates were not to be made available at the new data zones until August 2015, which would mean that calculation of the cost ratios and most of the candidate needs indicators could not be done at the new geography until then. This release would include mid-year estimates from 2001 to 2014, and future mid-year estimates would definitely be available at the new geography. On the other hand, mid-year population estimates at the old data zones – while they were currently available for years up to 2013 – were not planned to be made available from year 2014 onwards.

In terms of the candidate needs indicators (section 6.1), the only data source that was unavailable at the new geographies was the Scottish Index of Multiple Deprivation (SIMD). While doing the analysis at the new geographies would mean an 8-month delay to wait for the population data, inclusion of SIMD in the candidate indicators would have entailed a further year’s delay: SIMD 2012 had been provided at the old 2001 data zones; SIMD 2016 was to be released at the new data zones in August 2016.

The subgroup decided it was appropriate to exclude SIMD from the potential candidate variables rather than wait for its release at the new data zones in August 2016. Section 6.1.3 outlines the reasoning for this, which related more to statistical validity and stability than to the data availability issue.

The subgroup also agreed to delay the review work by 8 months in order to take advantage of the new geographies. This was reasoned to best fulfil the TAGRA core criteria: particularly in terms of relevance and responsiveness, in that the constructed model would be built upon an up-to-date population distribution; and in terms of stability, since the new geography would be unlikely to change for a further 10 years.

### 2.2 Data used in the review work

The analysis to look at the costing methodology (chapter 3) was done in 2014 at the old 2001 geographies:

* The cost ratios were calculated using 2012/13 inpatient and outpatient activity data and costs, with 2012 mid-year estimate populations.
* The 2011/12 Scottish National Tariff was used in Healthcare Resource Group (HRG) cost analyses.
* The 2012/13 Patient Level Information Costing System (PLICS) costed SMR01 file was used for PLICS analyses. The data used included inpatients only, excluded any patients with zero costs, and excluded patients treated at private hospitals.

The analysis to support the decisions on geographical granularity and time span for the cost ratios (section 4.2) was done in early 2015, at the 2001 geographies. This was done using the ‘reference model’ – that is, the current Acute MLC model as outlined in section 1.3. The data used was as follows:

* The 1-year cost ratios were calculated using 2012/13 inpatient and outpatient activity data and costs, with 2012 mid-year estimate populations.
* The 3-year cost ratios were calculated using 2011/12, 2012/13 and 2013/14 activity and costs, with 2011, 2012 and 2013 mid-year estimate populations.
* The All-cause SMR <75 variable was calculated using death records from 2008 to 2012 calendar years, with 2012 mid-year estimate populations.
* The LLTI variable was calculated using 2011 Census data for the first time, with 2011 mid-year estimate populations. There had been a change in the response options for the relevant Census question from a simple “Yes”/“No” to three options: “Yes – a little”, “Yes – a lot” and “No”. The options for updating the variable were investigated, and a variable using counts of both positive responses was adopted.
* The supply variables used were the ones currently in use in the NRAC formula, which were last recalculated in 2011 using 2009/10 hospital activity data.

The initial analysis to explore unmet need methodologies (chapter 8) was done in mid-2015, using the same data as for the geography and time span analysis above.

The supply model selection work (chapter 5), the indicator selection work (chapter 6), and the subsequent analysis on possible age splits (section 4.4), additional indicators (chapter 7) and unmet need (chapter 8) was done from August 2015 onwards at the new 2011 geographies:

* The 3-year cost ratios used in the regressions were calculated using 2011/12, 2012/13 and 2013/14 activity and costs, with 2011, 2012 and 2013 mid-year estimate populations. The case-mix adjustment was calculated using 2013/14 PLICS data.
* A 1-year “future” cost ratio was used to evaluate predictive power of the fitted models; this was calculated using newly available 2014/15 activity and costs, with 2014 mid-year estimate populations.
* The candidate supply variables were calculated at the new data zones using 2013/14 hospital activity data, GP counts as at 1st July 2015, 2011 data zone population centroids, and 2013 mid-year estimate populations.
* The data used for the candidate indicators is outlined fully in chapter 6.

### 2.3 Effect of the redrawn geographies

One of the first tasks carried out once the 2011 data zone populations became available was to rebuild the reference model at the new geography and compare the fitted model with that at the old 2001 geography.

The adjusted R2values were found to be broadly comparable, with no substantial changes. The coefficient of the needs index however increased significantly for several diagnostic groups. This suggested that the new geography had resulted in a more sensitive response: a bigger increase in cost can be inferred from a given increase in the needs index – possibly because the new geography has resulted in more socio-economic homogeneity within each data zone than before, allowing variation in need to be modelled more effectively. In general, there were slightly more cost ratio outliers at the new data zones; however, the percentage of *residual* outliers was the same or slightly *lower* than that at the old data zones, and was much lower than the percentage of cost ratio outliers.

## CHAPTER 3: COSTING METHODOLOGY

### 3.1 Current costing methodology

The MLC cost ratios (used as the dependent variable in the regression analysis) are the ratio of the actual costs of a small area to the expected costs based on the age and sex structure of the area’s population. The costing methodology refers to the method used to estimate these costs.

The Costs Book, published annually by ISD, is the only source of published costs information for NHSScotland. Cost data is collected at specialty, patient type, and hospital level; but not by patient demographics such as age and sex, or by location of patient residence. The Scottish Morbidity Records (SMRs) provide detailed information on episodes of care, including patient age, sex and data zone of residence, as well as diagnostic and treatment information including length of stay.

The current NRAC costing methodology combines the detailed SMR activity data with specialty-level cost data to estimate a cost for an individual episode of care. These can then be aggregated (a) over data zone of residence to produce the ‘actual’ costs, and (b) over age and sex categories at Scotland level in the calculation of expected costs. For the latter, the Scotland-level costs and populations are analysed within 20 age groups – 0-1, 2-4, 5-9, 10-14, etc, up to 85-89, 90 years and over – and then combined with the populations of data zones (in the same age-sex groupings) to produce expected costs.

For daycases, assigning a portion of the Costs Book cost to each episode is straightforward: the daycases cost for each specialty is simply divided equally between all episodes within the specialty. The same procedure is used for costing outpatient appointments.

For inpatients, a part of the inpatients cost is divided equally between episodes in the same way; this is referred to as the ‘fixed’ cost component. The remainder of the cost – the ‘variable’ cost component – is divided between episodes based on the length of stay, with each day of stay attracting an equal portion of the ‘variable’ cost.

The fixed cost component represents costs such as overheads, theatres, and medical staff – a “cost of admission”. The variable cost component represents direct costs such as nursing staff and pharmacy which are thought to be proportional to the length of stay of the patient – a “cost per day”.

The ratio used to split inpatient costs into the fixed and variable components depends on the specialty, and was derived in the NRAC review from a regression analysis of the relationship between average cost per episode and average length of stay (across hospitals) within each of five specialties: General Medicine, General Surgery, Gynaecology, Obstetrics, and Special Care Baby Unit. The results were then applied to all other specialties. In the regressions the observations were weighted by number of discharges to take into account different hospital sizes.

The alternative methods considered by the review for estimating costs for episodes of inpatient care are outlined in the following section. Papers TAMLC04, TAMLC05, TAMLC07 and TAMLC08 from subgroup meetings relate to the investigations into the costing methodology. The alternative methods differ from the current NRAC method in how they use the two basic sources of data – Costs Book costs and SMR activity – along with other information. One key distinction is in the ‘case-mix adjustment’, which is the means by which more complex episodes of care pick up a higher cost; in the NRAC method, this is generated by (1) the difference in average costs between specialties, (2) the difference in length of stay between episodes, and (3) the ratio used to split inpatient costs into fixed / variable components.

### 3.2 Options considered in the review

#### 3.2.1 Alternative costing methodology

The alternative costing methodology considered was the Patient-Level Information Costing System (PLICS), originally developed by NHS Highland and currently used by ISD’s Integrated Resource Framework (IRF) team as part of their work to support the integration of health and social care. It utilises the specialty cost components reported in the Costs Book rather than simply the net totals, which are in two broad categories: direct cost components are those which can be directly attributed to the particular clinical service or patient; allocated costs are those which cannot – including indirect costs (such as laundry), and overheads (such as building maintenance). The PLICS costing methodology apportions hospital- and specialty-specific direct costs to individual patient records, based on length of stay, any theatre time, and information about specific high cost items used e.g. prosthetics.

PLICS was still regarded as “developmental”, with no firm timescale for the completion of the developmental work. There were therefore concerns about its availability, accuracy, and stability. The subgroup therefore decided to retain the current NRAC costing method for the Acute MLC review, reviewing only the case-mix adjustment aspect of the methodology. In the longer term, however, PLICS costs should be considered for use within the NRAC formula as the PLICS costing method is more sophisticated than the current NRAC approach.

#### 3.2.2 Alternative case-mix adjustments

Having decided to retain the broad approach of the NRAC method of costing, the review considered alternative ways of calculating the fixed/variable ratios which are applied to the total specialty costs in the first step. The subgroup recognised the limitations to the current NRAC case-mix approach, which are a consequence of the fact that it was developed when only relatively limited costing data was available. The regressions were run for only 5 specialties, with the results then assumed to apply to all other specialties. The observations themselves, representing hospitals, were few in number and very highly aggregated.

Three alternative approaches to a case-mix adjustment were considered: two based on PLICS data, and one based on relative Healthcare Resource Group (HRG) costs.

The first PLICS-based method (the “PLICS components method”) would derive the cost of admission and cost per day percentages by separately summing the relevant PLICS cost components for each episode of care in every specialty. This would follow the way that the PLICS costing method itself distributes the direct cost components between episodes – some being apportioned as a fixed cost related to an admission, others being apportioned based on the length of stay. ‘Direct’ costs alone would be used; PLICS assigns allocated costs to episodes in proportion to the fixed and variable components of the direct costs, such that their inclusion would make no difference to the percentage cost splits.

The costs applied per day (i.e. “per day” costs) consist of the following direct costs:

* Medical
* Nursing
* Laboratory
* Radiology
* Pharmacy
* Allied Health Professional Other
* Other direct care (after any High Cost Items exclusions)

The costs applied on admission include the following direct costs:

* Medical
* Laboratory
* Radiology direct costs

High cost items costs and theatre department / medical procedure related costs are considered to be “admission” type costs.

The second PLICS method (the “PLICS regression method”) returns to the linear regression approach to deriving case-mix adjustment parameters, as used in the current NRAC method. However, instead of using the *average* cost per episode and length of stay (by hospital) as the regression variables, the PLICS patient-level costing allows for the use of *individual* cost per episode and length of stay instead.

The third alternative case-mix approach would use Healthcare Resource Groups (HRGs), a case mix classification used by the NHS in England and maintained by the National Casemix Office (NCO) at the Health & Social Care Information Centre (HSCIC). HRGs organise patient activity into clinically meaningful groups with similar resource intensity, based mainly on diagnosis and procedure information (ICD-10 and OPCS-4 codes). Scottish HRG-based costs are already calculated by ISD for the Scottish National Tariff; specialty costs from the Costs Book are disaggregated by HRG based on the assumption that the relative cost difference between any two procedures or conditions in Scotland is the same as in England.

In the third case-mix adjustment method, the portion of cost applied to each unit of activity would depend on its HRG code rather than the length of stay. There would no longer be a simple ‘fixed’ / ‘variable’ division. A further change would be that the basic activity unit would have to be a spell within specialty, rather than an episode.

Analysis was carried out to compare the cost ratios resulting from the different case-mix adjustment methods. All four approaches yielded generally similar cost ratios. The subgroup decided against the method using HRG costs, for the following reasons:

* The HRG method was found to lead to larger differences from the current method than either of the PLICS methods; the PLICS methods therefore offered better stability.
* Statistical test results did not provide evidence that HRG would form a better case-mix adjustment variable than length of stay.
* The HRG codes for which Scottish costs were available were somewhat out of date, and there was doubt over whether the Tariff would be updated going forward.
* The HRG costs, being based on weights derived from English HRG costs, are less sensitive to length of stay since long stays are truncated in the English HRG cost calculations.
* The unit of activity for the HRG costs (spells within specialty) was also different from the episode-based activity measure currently used in NRAC.

The PLICS methods were regarded as offering a significant improvement over the current NRAC case-mix adjustment. The majority of specialties where the NRAC fixed and variable ratios differed substantially from the PLICS-based ones occurred in cases where the NRAC results had been extrapolated from another specialty. With continued development of the wider PLICS methodology over the coming years, the adoption of a PLICS-based option would represent a forward-looking choice.

After extensive investigation and discussion, the “PLICS components” method was ultimately recommended over the “PLICS regression” method for the case-mix adjustment, for the following reasons:

* It will be more practical to calculate for all specialties as it is a simpler process.
* It was regarded as more transparent and easy to explain.
* For Rheumatology, the use of the “PLICS components” method would avoid the 0% cost of admission result suggested by the “PLICS regression” method.
* The use of a simple regression-based method is not statistically well justified in the case of specialties for which substantial differences in costs per day exist between hospitals. In these cases the components derived case-mix adjustment parameters are likely to be more robust.
* The PLICS methodology applies costs to episodes based on an assumed relationship between direct cost components and length of stay. There is therefore a risk of over-estimating model-fit / effect-size in the “PLICS regression” case-mix adjustment, through non-independence of length of stay and PLICS cost.
* The PLICS methodology allows different relationships to exist between direct cost components and length of stay, between different hospitals. The formal statistical requirements of the regression method cannot be met in these cases.

Recommendation 1: retain the NRAC costing method, but with the case-mix adjustment derived from summing PLICS components by specialty

## CHAPTER 4: MODEL SPECIFICATION

### 4.1 Functional form

Paper TAMLC10 presented the investigations relating to the functional form of the MLC regression model. The subgroup decided upon a linear modelling approach, without transformations, for the following reasons:

* Predictive power: research by the subgroup, and previous work carried out in the NRAC review and the Mental Health and Learning Difficulties MLC review, indicated that a linear approach performed equally well in terms of predictive power compared to alternative modelling approaches.
* Simplicity and transparency: in line with TAGRA’s core criteria, the linear approach without transformations was deemed more transparent, allowing for the ability to trace through the implications of variation in the needs index to variation in the predicted value.
* Resource constraints: comparing alternative functional forms is difficult and therefore time-consuming. Even if a specific functional form is judged better with a particular set of variables, a change in the variables would require the comparison of functional form to be made again.

Recommendation 2: retain the linear functional form

### 4.2 Geography basis and time span for the cost ratios

In the NRAC 2007 review, 1 year of data was used to calculate the cost ratios at intermediate zones, for all care programmes, since more than one year of costed data was not available. In terms of geography, intermediate zones were used for all care programmes. Data zones were rejected because of high levels of zero activity for some condition groups (Mental Health and Learning Difficulties in particular), and higher year-to-year variations in costs and numbers of patients. On the other hand, there was a concern in using intermediate zones that the effect of small “pockets of deprivation” would be lost or smoothed out.

The Mental Health and Learning Difficulties (MH & LD) MLC review resulted in a move to using 3 years of data for the MH & LD cost ratios, with the geography basis unchanged.

Paper TAMLC22 reported the analysis to support decisions on the geography and time span for the Acute MLC cost ratios. It was decided that using 3 years of data for the cost ratios was preferable to using just a single year. Using 3 years would smooth out random year-to-year variations which would be likely to improve stability, while remaining sufficiently responsive to changes over time.

The subgroup considered at length the question of geography – whether to keep the MLC adjustment at intermediate zones or instead move to data zones. Analysis was carried out using the reference model at the 2001 geographies to support the decision. The main discussion points were as follows:

* Aggregation: In principle, using the lowest possible level of disaggregation is usually advisable for analysis whenever data is available. Moreover, unlike for the other care programmes, aggregating data is not necessary for the MLC adjustment for Acute services since there is no zero activity at data zones.
* Linearity and ‘noise’: A weaker relationship between the cost ratios and the needs index was observed at data zone level. One of the linear model assumptions – normality of the model residuals – was violated to a greater extent with data zones, but it was agreed that this was not an issue for the linear regression because of the large sample size (6,976 data zones in the 2011 geography). Furthermore, the analysis showed that there were no influential points changing the Acute MLC index coefficient, and the needs index was a significant predictor of the cost ratios. The main consequence of the weaker relationship and the greater amount of ‘noise’ was that the explanatory power (adjusted R2) of the data zones’ model was lower. Aggregation of model predictions from data zone to intermediate zone level allowed direct comparison of predictive power and this reassured the subgroup that on average, over the whole range of values, neither approach would give consistently better predictions than the other at the intermediate zones level. Some members argued that it would be more difficult to find indicators to explain the variation in cost, based on the noisier data zone cost ratios; others disagreed.
* Regression coefficients: The needs index coefficient for all diagnostic groups was found to be systematically higher at data zone level, which was considered by some members to imply better capturing of pockets of deprivation. This is in line with the equity and responsiveness elements of the TAGRA core criteria. However, others interpreted the difference in coefficient simply as a reflection of the decreased variance of the cost ratios when aggregating to intermediate zones.

There was no outright consensus among all members. However, core criteria were reviewed and a majority agreed that data zones would be the better choice for the geography basis of the Acute MLC adjustment. The recommendation is based mainly on the TAGRA core criteria, since finer granularity should lead in principle to better equity and responsiveness, and on the absence of a convincing reason to continue aggregating Acute data to larger geographical units.

Recommendation 3: change the time span of the Acute MLC cost ratios from 1 year to 3 years

Recommendation 4: change the geography basis of the Acute MLC adjustment from intermediate zones to data zones

### 4.3 Diagnostic grouping

The current Acute MLC adjustment analyses cost variation within subsets of the Acute activity, referred to as diagnostic groups. These are: Cancer, Heart, Respiratory, Digestive, Injury, Outpatients and Other. Separate linear models are fitted for each of these diagnostic groups, although a common needs index is used for all. Five of the current diagnostic groups – Cancer, Heart, Respiratory, Digestive, and Injury – correspond to high-level ICD-10 code groups [5], and activity is selected into each diagnostic group based on its main ICD-10 code. A sixth group, Other, contains all other inpatient and daycase activity. The seventh group, Outpatients, contains all Acute outpatient activity (relating to any Acute condition). See section 1.1 for a summary of the comparative activity and costs for each of these diagnostic groups.

Advice was sought from an ISD consultant in public health on whether any other structure might be better. The feedback was that ICD-10 remains a suitable way of classifying activity into high-level diagnostic groups, but that Outpatients stood out as being a *patient category* rather than a diagnostic group. Ideally it would be better to calculate diagnostic group cost ratios that combined all inpatient, daycase and outpatient costs for the condition types; however, this was not possible due to the lack of ICD-10 coding on outpatient activity. The subgroup discussed the possible use of specialty information instead to assign diagnostic groups to Outpatients activity; however, it appeared that this would result in the majority of activity (82%) being assigned to ‘Other’ which was already a large group. For these reasons, it was decided to keep the Outpatients group separate.

The subgroup considered the merits of further disaggregation. ‘Other’ is a large diagnostic group containing the Acute inpatient and daycase activity falling within all the other ICD-10 code groups besides Cancer, Heart, Respiratory, Digestive, and Injury. The contents were examined (Table 2) to see whether any other code groups could be drawn out from Other to become Acute diagnostic groups in their own right. The only code group within Other that was comparable in size to the existing diagnostic groups was ‘Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified’, which did not seem a good candidate for a separate diagnostic group.

On the other hand, re-aggregation was also considered: the option not to analyse cost variation in diagnostic groups at all, but instead, to develop a single linear model for the Acute care programme as a whole – a ‘Whole Acute’ option. This option was kept open throughout the indicator selection process; Whole Acute results were included in all analyses, alongside results for the diagnostic groups.

Although the coefficient of the candidate needs indices varied between diagnostic groups, there proved to be no significant differences between diagnostic groups in terms of which indicators were selected (see Chapter 6). The final choice was therefore between retaining diagnostic groups with a common needs index (but with different coefficients for each diagnostic group), or changing to a ‘Whole Acute’ model.

To compare the model performance of the ‘Whole Acute’ option with that of the diagnostic groups, analysis was carried out, at a late stage of the indicator selection process, using the candidate models (see paper TAMLC40). Predictions of cost from the separate diagnostic groups were aggregated together, by the same method used routinely in the NRAC shares calculation, and compared with the predictions of cost from a ‘Whole Acute’ regression. The residual sum of squares (RSS) was computed, based on comparing the predictions from both methods with a 2014/15 cost ratio – a “future observation”. The predictive power was virtually the same regardless of whether diagnostic groups were used or not. However, the geographical distribution of activity within the diagnostic groups may not be constant, and while modelling their costs separately does not improve the predictions overall, this may not be true for smaller regions.

Ultimately, the subgroup preferred to retain diagnostic groups on grounds of face validity, possible improvement to the predictions at local level, and relevance to a range of stakeholders.

Recommendation 5: retain the Acute diagnostic groups

Table 2: Acute expenditure and activity within ‘Other’, by high-level ICD10 classification

|  |  |  |
| --- | --- | --- |
| **ICD-10 group** | **Actual spend 2013/14** | **Number of episodes 2013/14** |
| Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified | £319m | 204,912 |
| Diseases of the musculoskeletal system and connective tissue | £221m | 92,259 |
| Diseases of the genito-urinary system | £198m | 101,434 |
| Diseases of the nervous system | £92m | 38,272 |
| Certain infectious and parasitic diseases | £87m | 40,811 |
| Factors influencing health status and contact with health services | £81m | 57,698 |
| Diseases of the skin and subcutaneous tissue | £71m | 32,433 |
| (missing) | £70m | 25,176 |
| Diseases of the eye and adnexa | £70m | 50,347 |
| Endocrine, nutritional and metabolic diseases | £51m | 28,893 |
| Mental and behavioural disorders | £38m | 16,406 |
| Congenital malformations, deformations and chromosomal abnormalities | £30m | 9,669 |
| Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism | £28m | 21,625 |
| Diseases of the ear and mastoid process | £9m | 6,653 |
| Certain conditions originating in the perinatal period | £2m | 1,281 |

### 4.4 Age split

The Mental Health and Learning Difficulties MLC review resulted in implementation of an age split (<65 and 65+), whereby separate cost ratios are calculated for the older and younger populations, and different needs indices are used to model additional need for the two groups. This was motivated by *a priori* evidence of different need drivers in different age groups. The subgroup considered whether the Acute MLC adjustment could be improved by introducing an age split.

Advice from ISD clinical consultants was that there were no confident clinical grounds on which to suggest a split by age group, and that the introduction of any such split should be driven by the data itself. Various concerns were raised within the subgroup about the idea of introducing an age split: the need to continually re-check its validity as other changes are made to the data or the model, in the absence of an *a priori* clinical steer; possible bias in trying to capture the interaction between age and MLC in a simple age split, when the interaction may vary geographically or between diagnostic groups, and may change with time; difficulty in explaining why age needs to be considered after having accounted for age and sex in the Age-Sex component of the formula. The subgroup decided that the indicator selection process would be done based on an all-ages model, and then subsequent analysis would be carried out using several trial age splits to check the performance of the model across age groups.

Paper TAMLC51 presented the analysis of possible age splits. Lower adjusted R2 and lower coefficient values were generally observed for the older age groups as compared to the younger groupings (see Table 3 for the coefficients). This indicated that there was less of a relationship between the needs index and cost for older populations; possibly because older age is already a strong predictor of cost (accounted for in the Age-Sex component of the formula) and additional needs become less important at older ages. Another factor is that in many cases, the proportion of activity accounted for by the older age group is relatively low; despite the steep increase in activity rates with age, the population also decreases with age. Lower activity levels will mean that there is a larger variance in the cost ratios and the model will explain less of the variation – this is also seen, for example, in adjusted R2 differences between large diagnostic groups (such as Outpatients and Other) and smaller ones.

The residual sum of squares (RSS) was used to compare the model predictions with the 2014/15 cost ratios: by aggregating the predictions for the two separate age groups, we can directly compare the overall predictive power of an age-split model with that of an all-ages model. The resulting RSS values are given in Table 4. Low RSS values indicate that the observations are relatively close to the predictions. It is clear that at the “all ages” level, there is very little difference between the predictive power of models with or without an age split. Where the RSS does differ slightly, it favours the model without an age split.

The subgroup was therefore content to retain the Acute MLC all-ages model.

Recommendation 6: continue to model all ages together for the Acute MLC adjustment

Table 3: Acute index regression coefficient values obtained from fitting the proposed index model, along with 95% confidence intervals in brackets, by diagnostic group and age grouping.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Diagnostic group** | **All-ages coefficient** | **<65 coefficient** | **<70 coefficient** | **<75 coefficient** |
| **≥65 coefficient** | **≥70 coefficient** | **≥75 coefficient** |
| **Cancer** | 0.045  (0.040, 0.500) | 0.063  (0.054, 0.072) | 0.063  (0.055, 0.071) | 0.062  (0.055, 0.069) |
| 0.046  (0.038, 0.054) | 0.043  (0.033, 0.052) | 0.031  (0.019, 0.043) |
| **Heart** | 0.094  (0.089, 0.099) | 0.188  (0.178, 0.197) | 0.175  (0.167, 0.183) | 0.163  (0.156, 0.170) |
| 0.071  (0.063, 0.079) | 0.058  (0.050, 0.067) | 0.042  (0.032, 0.053) |
| **Digestive** | 0.101  (0.097, 0.105) | 0.141  (0.135, 0.147) | 0.140  (0.134, 0.146) | 0.135  (0.129, 0.140) |
| 0.083  (0.075, 0.091) | 0.065  (0.056, 0.074) | 0.053  (0.042, 0.065) |
| **Injury** | 0.097  (0.093, 0.102) | 0.172  (0.165, 0.179) | 0.171  (0.164, 0.178) | 0.165  (0.159, 0.171) |
| 0.062  (0.053, 0.071) | 0.045  (0.034, 0.055) | 0.037  (0.025, 0.049) |
| **Other** | 0.084  (0.081, 0.086) | 0.127  (0.122, 0.131) | 0.125  (0.121, 0.129) | 0.121  (0.117, 0.124) |
| 0.063  (0.058, 0.067) | 0.053  (0.047, 0.058) | 0.042  (0.036, 0.049) |
| **Respiratory** | 0.154  (0.149, 0.160) | 0.200  (0.190, 0.210) | 0.213  (0.204, 0.222) | 0.217  (0.209, 0.225) |
| 0.179  (0.170, 0.188) | 0.161  (0.151, 0.171) | 0.136  (0.125, 0.148) |
| **Outpatients** | 0.035  (0.033, 0.037) | 0.046  (0.044, 0.049) | 0.045  (0.043, 0.047) | 0.043  (0.041, 0.045) |
| 0.005  (0.002, 0.009) | -0.004  (-0.008, 0.000) | -0.016  (-0.022, -0.010) |

Table 4: RSS values obtained from comparing the proposed needs index model with the 14/15 cost ratios, for various possible age splits, after aggregating the age-group predictions.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Diagnostic group** | **RSS – model with no split** | **RSS – model with split at 65** | **RSS – model with split at 70** | **RSS – model with split at 75** |
| **Cancer** | 3264 | 3264 | 3263 | 3263 |
| **Heart** | 3618 | 3621 | 3621 | 3620 |
| **Digestive** | 2200 | 2204 | 2205 | 2203 |
| **Injury** | 2852 | 2854 | 2855 | 2853 |
| **Other** | 828 | 831 | 831 | 830 |
| **Respiratory** | 3569 | 3574 | 3579 | 3582 |
| **Outpatients** | 396 | 396 | 396 | 395 |

## CHAPTER 5: SUPPLY MODEL

### 5.1 Rationale for the supply model

The purpose of the MLC adjustment is to predict the geographic variation in healthcare costs that is due to variation in *need*, by analysing the relationship between cost ratios and the indicators of need. In order to be confident of the validity of the inferred relationship, any other factors influencing the cost ratios should be taken into account where possible. One such factor is the likelihood of higher healthcare activity rates in places where healthcare is more readily available – either because need is not being met adequately in areas with lower supply, or because need is being “over-met” in areas with high supply. In particular, there is a need to control for the effect of previous budget allocations on the pattern of supply of hospital services.

For this reason, supply variables always have been included in the NRAC formula’s MLC regression model. The supply variables are control variables: that is, they are included in the regressions to ensure unbiased estimates of the coefficients representing the relationship between cost and the indicators of need, but are not used in the *prediction* of cost. To include them in the prediction of cost would disadvantage any areas in which the level of healthcare provision was lower than it needed to be – or, conversely, reward areas in which higher levels of activity resulted purely from greater supply. Including these variables in the *regression* assures us that the effect of supply is accounted for, and we can be more confident that the resulting model uses the indicators of need to predict only cost variations that are due to differences in need, and not due to variations in supply.

### 5.2 Candidate supply variables

Table 5 shows the candidate supply variables that were previously explored in the 2007 NRAC review and have now been re-examined in the current Acute MLC review. Waiting times were considered for inclusion, but the subgroup agreed that waiting times information was too complex to be suitable for use, with inconsistency in the recording of information across NHS Boards.

The current supply model, for all care programmes, consists of the health board dummies along with IPACX and OPACX.

Table 5: Potential supply variables for investigation

|  |  |
| --- | --- |
| **Variable** | **Description** |
| GPCount | Number of whole time equivalent (WTE)[[1]](#footnote-1) GPs serving each data zone. This variable is the sum of the number of practices serving each data zone weighted by the whole time equivalent number of GPs in each practice. |
| GPSup, GPSup5, GPSup10 | A function of both the number of whole time equivalent (WTE) GPs serving each data zone and the distance to the practices. This can be expressed algebraically as:  where *n* (the “intrazonal cost”) is 1, 5 or 10 kilometres. The intrazonal cost is an assumption about the part of the travel *within* the small area; the distance to practice represents travel *between* the small area (i.e. its population centroid) and the practice. |
| IA1, OA1 | Size of nearest inpatient / outpatient facility, where size is measured as the number of inpatient or outpatient episodes in the last year. Nearest here means closest to the data zone population centroid. |
| IA2, OA2 | Size of nearest or second-nearest facility, whichever is the larger. |
| IPACX, OPACX | A function of both the size of the inpatient / outpatient facilities serving each data zone and the distance to the facility. This can be expressed algebraically as: |
| IPAC, OPAC | Similar to IPACX and OPACX, but with an attempt to correct the size of the facility for the size of population that it serves. Expressed algebraically as:  where the summation over ‘h’ is a summation over all hospitals serving the given data zone, and the summation over ‘d’ is over all data zones served by the particular hospital. |
| Health Board dummy variables | Represents the effect of the health board. There is one binary variable representing each Health Board; it takes the value ‘1’ for data zones within the Health Board and ‘0’ otherwise. |

The supply variables form three conceptual groups: GP supply, hospital supply, and health board dummy variables. The hospital and health board variables relate clearly to supply of secondary care; however, the role of the GP supply variables is more ambiguous. GP supply could be seen as either a substitute for secondary care or as a complement. GPs are substitutes when they treat patients who would otherwise have gone to A&E or would have needed more serious and higher cost treatment without the early intervention provided by more readily available primary care. In contrast a higher supply of GPs might generate additional episodes of hospital care through a higher level of referrals.

The health board dummies aim to capture broad differences between the 14 NHS Boards. The GP and hospital supply variables are all intended to represent the availability of healthcare at a finer-grained level, in various ways. The simplest of these (GPCount, IA1, OA1, IA2, OA2) look only at the ‘size’ of the facilities serving each data zone – in terms of numbers of GPs, or episode counts for nearby hospitals – with the rationale being that bigger facilities represent a greater availability of healthcare (or, perhaps, a perception thereof on the part of patients). More complex variables (GPSup, GPSup5, GPSup10, IPACX, OPACX) include a term for the distance between the data zone and the facility, since facilities that are nearby are more accessible than facilities some distance away. Finally, the most complex variables, IPAC and OPAC, apply a further correction that accounts for the size of the population served by each facility. This can be thought of as an adjustment for “demand”. A large hospital can treat many patients, but if a very large number of potential patients live close by then this might not amount to a high availability of healthcare (or a perception thereof) for an individual. IPAC and OPAC try to account for this [6].

### 5.3 Selection of supply variables

Paper TAMLC39 presented analysis to support selection of the supply model. The methodology of the 2007 NRAC review treated the hospital supply variables as pairs, with the inpatient and outpatient components for each variant taken together. The performance of the different supply variables (or pairs of variables), when entered into the regression model along with the current needs index and health board dummies, was examined by looking at adjusted R2. The best-performing supply variables were IPAC / OPAC and IPACX / OPACX, with the adjusted R2 being virtually identical for these two pairs.

However, since most hospitals are both inpatient and outpatient facilities, these variables are highly correlated, and so the subgroup considered it more reasonable to use only one of the pair – OPAC or OPACX for Outpatients, IPAC or IPACX for all other diagnostic groups. Indeed, there was no noticeable difference in adjusted R2 when using just a single variable compared to the pair.

Including a supply model in the linear regression has an effect on the coefficient of the Acute needs index, since some of the variation in costs is now explained by the supply variables. Table 6 shows the effect, for each of the considered supply models. Generally the effect of including the supply models is to lower the needs index coefficient, although this is not the case for the diagnostic group Injury. There is almost no difference between the considered supply models in terms of the resulting needs index coefficients.

Table 6: Acute needs index coefficient values obtained from regression using different models

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | **Whole Acute** | **Cancer** | **Heart** | **Digestive** | **Injury** | **Other** | **Respiratory** | **Outpatients** |
| needs index only | 0.097 | 0.063 | 0.113 | 0.131 | 0.113 | 0.104 | 0.197 | 0.047 |
| health board dummies + needs index | 0.091 | 0.055 | 0.117 | 0.117 | 0.121 | 0.098 | 0.187 | 0.037 |
| IPAC + OPAC +  health board dummies + needs index | 0.090 | 0.054 | 0.115 | 0.116 | 0.120 | 0.096 | 0.185 | 0.037 |
| IPACX + OPACX +  health board dummies + needs index | 0.089 | 0.054 | 0.115 | 0.116 | 0.120 | 0.097 | 0.184 | 0.037 |
| IPACX +  health board dummies + needs index | 0.090 | 0.054 | 0.115 | 0.117 | 0.121 | 0.097 | 0.184 | - |
| OPACX +  health board dummies + needs index | 0.090 | - | - | - | - | - | - | 0.036 |

Technical Report D of the 2007 NRAC review expressed some concerns around IPAC and OPAC, noting that some remote data zones in the Island boards had similar values of IPAC / OPAC to some very urban data zones. We find that this difference between IPAC / OPAC and IPACX / OPACX still holds: for example, for data zones in NHS Greater Glasgow & Clyde, OPACX values range from 227 to 2,134, giving no overlap with the range of 4 to 44 for NHS Orkney; OPAC by contrast ranges from 0.11 to 0.54 for NHS Greater Glasgow & Clyde and from 0.02 to 0.47 for NHS Orkney, so there is a significant overlap in values between these two boards.

What this shows is that when using IPAC and OPAC, distant hospitals that serve only very small populations (as in Orkney) have a similar weight in terms of “supply” to hospitals close by that serve much larger populations (as in Glasgow). By contrast, IPACX and OPACX, which do not consider the population size, regard urban areas – where the distances to hospitals are much shorter – as having much greater levels of supply.

IPACX and OPACX have higher correlations with the cost ratios than IPAC and OPAC, and result in slightly higher adjusted R2 in a few diagnostic groups. They also have a simpler, more easily explainable formulation, and are currently in use within the NRAC formula.

However, it makes no practical difference to the result which of these options is used. This perhaps simply reflects that the supply variables (whichever ones are used) only explain a very small proportion of the variation in cost. Indeed, the increase in adjusted R2 obtained when including the supply variables – as compared to a regression on just the health board dummies and needs index – is less than 1 percentage point for all diagnostic groups and typically 0.1—0.2 percentage points.

Other possibilities were briefly explored to see if the contribution of the supply variables could be improved. Firstly, the “buffer distance” of 10km for IPAC / OPAC and IPACX / OPACX (see equations in Table 5) was varied, as it had been for the GPSup variables, and an option to make the buffer distance dependent on the geographical area of the data zone was also tested. Secondly, while the inverse *square* of the distance is used in the equations, this is somewhat arbitrary and so an inverse cubic function was also tried. None of these variants showed any real difference in the results.

The subgroup agreed to continue to use IPACX and OPACX as the supply variables, along with the health board dummies: IPACX for inpatient diagnostic groups and OPACX for Outpatients. (For the ‘Whole Acute’ option, IPACX was used, since inpatient data makes up the majority of this group.) The reasoning was based on the TAGRA core criteria:

* Practicality: IPACX and OPACX are easier to calculate and are currently in use within the NRAC formula.
* Transparency and face validity: IPACX / OPACX have a simpler, more easily explainable formulation; similar values of supply across diverse urban-rural settings as seen with IPAC / OPAC would be harder to justify.
* Objectivity: the analysis did not strongly favour either option, but IPACX and OPACX did have higher correlations with the cost ratios, and slightly higher adjusted R2 in a few diagnostic groups.

Recommendation 7: continue to use IPACX and OPACX as the supply variables, along with the health board dummies, but use *only* IPACX for inpatient diagnostic groups and *only* OPACX for Outpatients.

## CHAPTER 6: NEEDS INDEX DEVELOPMENT

### 6.1 Candidate variables

From February 2014 to December 2015, the subgroup discussed potential candidate variables for testing and investigated data availability. (Details of these discussions can be found in papers TAMLC3, TAMLC9, TAMLC20, TALMC23, TAMLC29, TAMLC34 and TAMLC36.) The list of previously-tested indicators formed the starting point, with additional variables included based on the subgroup’s knowledge of new data sources and of theoretical associations. The subgroup agreed to consider only indicators which have a theoretical link to Acute health need (for transparency and face validity reasons), that have a geographical granularity at least as low as intermediate zones, for which the available data is no more than 10 years old, and that are able to be updated going forward.

The following sub-sections outline the data sources explored for potential candidate variables, and the decisions made (outside of the selection process) in each case. Table 7 gives the full list of indicators that were entered into the selection process, with details of the construction of each variable.

#### 6.1.1 Scottish Patients at Risk of Readmission and Admission (SPARRA)

SPARRA is a risk prediction tool developed by ISD within the Integrated Resource Framework (IRF) to support health and social care integration. It predicts an individual's risk of being admitted to hospital as an emergency inpatient within the next year.

On investigating using the SPARRA data, the subgroup concluded it was unsuitable as a candidate indicator as it takes the form of a score (known as a SPARRA score) for each patient. Many of the risk factors and variables used in SPARRA were already considered in the review or were highly correlated with other variables considered.

#### 6.1.2 Quality and Outcomes Framework (QOF)

The Quality and Outcomes Framework (QOF) data on prevalence of various conditions was considered, but this is currently only available at GP Practice level. In the longer term, data from SPIRE (Scottish Primary Care Information Resource) – which is currently in development – may provide useful information at individual patient level; this could be considered in future MLC reviews.

#### 6.1.3 Scottish Index of Multiple Deprivation (SIMD)

Having decided to use the new 2011 geography for the Acute MLC review, inclusion of SIMD in the candidate indicators would have entailed a further delay: SIMD 2012 had been provided at the old 2001 data zones; SIMD 2016 was to be released at the new data zones in August 2016.

The subgroup discussed whether SIMD variables were appropriate as potential indicators of need, noting that, for transparency and statistical validity reasons, the original NRAC review had concluded that they were not. Concerns were highlighted around the use of the SIMD’s composite scores as predictor variables, given the non-linear transformations that occurred in the calculation of some domains, for subsequent use in the Acute needs index. There were also concerns about stability: the way the domain scores are composed is out of our control and may change in future SIMD updates; in particular, those domains with rank-based non-linear transforms (Health, Education, Access) will potentially be very sensitive to even minor changes in compositing method.

Some exploratory analysis was carried out at the 2001 data zones on the potential explanatory power of SIMD, based on SIMD 2012. The results suggested that the SIMD income domain score was as powerful a predictor of need as either of the current Acute indicators, LLTI or All-cause SMR <75. The addition of SIMD to the existing model however would add only a modest (generally less than 2% of cost ratio variance) amount of *additional* explanatory power.

The subgroup decided it was appropriate to exclude SIMD from the potential candidate variables, given the other candidate variables in the list, rather than wait for its release at the new data zones in August 2016.

#### 6.1.4 NRS death records

Following previous reviews, standardised mortality ratios and death rates were calculated using NRS death records. A cut-off age was used so that these variables represented “premature” mortality. Various causes of death and various cut-off ages were utilised to construct a range of potential indicators for testing.

#### 6.1.5 Census-based morbidity data

All 2011 census questions relating to health were utilised, resulting in variables representing the prevalence of:

* Long-term illness
* Mental health conditions
* Limiting long-term illness
* People who are long-term sick and not seeking work
* Poor general health

The Mental health condition data was included because of a theoretical link between mental health conditions and need for Acute services [7].

#### 6.1.6 Census-based deprivation data

Counts of job seekers, and of respondents with low education level, were used.

#### 6.1.7 Unpaid care

Carers were identified by the subgroup as being a high profile group in current policy. Consideration of carers is relevant for their own healthcare needs and for the costs to the NHS if they become unable to care. The impact of caring on use of Acute services is not currently known, but it was suggested that this could be tested in the model.

The 2011 census contained a question on unpaid care (“*Do you look after, or give any help or support to family members, friends, neighbours or others because of either: long-term physical / mental ill-health / disability; or problems related to old age? Do not count anything you do as part of your paid   employment*) with response options to indicate various numbers of hours per week. Variables constructed from the response data were entered into the selection process.

In the final stage of the process, the subgroup discussed concerns about the transparency and face validity of the provision of unpaid care variable. Clinical advice was sought and two published reports informed the discussion [8, 9]. The concerns discussed were as follows:

* The mechanisms through which unpaid care provision is associated with healthcare costs are not well understood.
* The data would only be updated every 10 years with a new census, and it is questionable whether geographical patterns of unpaid care would remain the same over that period, especially given the ongoing work to improve social care.
* The reports indicate that we should see a dose response relationship between unpaid care provision and higher Acute health costs, which was not observed in the modelling.
* It is difficult to infer causation in the relationship between unpaid care provision and utilisation; provision of unpaid care may be a proxy for higher health needs in a population, reflecting a higher number of people *requiring* care. Unpaid care may have a relationship with health costs in both directions, in that it does reflect a population with higher health needs but may also decrease demand for unscheduled care and reduce delayed discharges through the support that carers provide.

For these reasons, unpaid care variables were excluded from consideration at that stage (see paper TAMLC49). The subgroup recommends that unpaid care is re-examined in future reviews if better ways of capturing the need responsively can be identified.

#### 6.1.8 Older people living alone

In addition to considering need related to carers, it was also noted that the absence of a carer could have an impact on Acute services, for example in exacerbating delayed discharges. The census data allowed numbers of older people living alone to be counted, using various cut-off ages.

#### 6.1.9 Prescribing data

Detailed data on prescriptions is collected by ISD at patient level. An ISD pharmacy advisor was consulted for advice on using this data to form potential indicator variables for Acute services. The advice was as follows:

* Diabetes prescriptions might be a good indicator, because there would be an increased likelihood of needing Acute services in patients for whom diabetes is medicated.
* A variety of drugs are prescribed for different respiratory conditions, so it would be difficult to isolate particular conditions (e.g. asthma, COPD) through pharmacy data, but a general ‘respiratory’ category could be useful.
* In terms of capturing the needs of the elderly population, dementia drugs were suggested as one possibility.
* Using drug counts as a proxy for mental health conditions was more difficult, as drugs such as antidepressants can be routinely prescribed for other conditions besides mental health. Methadone was raised as a possibility, but it turns out to have a low CHI capture rate and therefore cannot be adequately mapped to data zone geography.

None of the three prescribing variables constructed had a high correlation with utilisation. The Diabetes and Respiratory prescribing variables had large numbers of zeros at data zone level; only the Dementia prescribing variable was therefore retained past the first stage of variable selection. Analysis at that point showed that when entered into a regression, this variable had a fairly high but counter-intuitively negative coefficient. The adjusted R2 achieved using this variable was relatively low, and in a scatter plot, no real relationship with the cost ratios was apparent. It was felt that this variable was not able to represent dementia prevalence adequately since not all dementia patients – and particularly not those with the most advanced dementia – would be prescribed medication. It was therefore excluded at that stage.

#### 6.1.10 High Resource Individuals (HRIs)

HRIs are now identified in ISD as part of the patient-level costing work carried out within the Integrated Resource Framework (IRF) to support health and social care integration. They are defined as follows:

1. Calculate the total resource for each individual service user during a financial year.
2. Rank in order of total resource consumption.
3. Identify each of the highest resource users as an HRI until cumulative expenditure reaches 50% of total expenditure.

The HRI variable was included in the final list of potential indicators, and proved to be a strong indicator of utilisation; however, the subgroup decided to exclude it due to concerns that as a “supply-side” variable it is not independent of past costs i.e. of the dependent variable in the model.

#### 6.1.11 Ethnicity

Ethnicity data from the 2011 census was obtained. Initially the Ethnicity variable had been planned to measure the fraction of the population that was non-white. However, experts on health inequalities present at the 10th meeting advised that a count of non-white people as a potential explanatory variable for healthcare use was problematic: it implied that all non-white groups were anticipated to share a common pattern of healthcare needs that differed from those of white populations. This was highly questionable when such a large number of different ethnicities are included. An independent Scottish Government report [10] indeed indicated that there are wide differences in health outcomes between different ethnic groups in Scotland. A decision was therefore made to use population counts of specific ethnic groups, using the above report as a guide to which ethnic groups should be grouped together. The following variables were constructed:

* Ethnicity I: all minority populations (including white minorities)
* Ethnicity II: all populations with worse than average health (Gypsy/Traveller and Pakistani).
* Ethnicity III: Gypsy/Traveller population
* Ethnicity IV: Pakistani population
* Ethnicity V: all populations with better than average health

The Ethnicity I variable, counting all minority populations, was not expected to perform well given the inequalities experts’ comments, but was included for continuity with past investigations with the intention to consult inequalities experts if it turned out to perform well.

The inclusion of Ethnicity V, counting all populations with better than average health, was motivated by the observation that some of the populations that would be included (e.g. White Polish) are much larger than the populations with poorer than average health; a variable using these populations could therefore be a better differentiator of need than Ethnicity variables II, III or IV. However, it was subsequently decided that this variable should not be used, because its relationship with the cost ratios would be in the opposite direction to that of the other indicators: higher values would be associated with lower cost. The subgroup had concerns about the face validity of appearing to penalise areas with more healthy populations, and about the practicality of combining variables into an index that had opposite-signed regression coefficients when entered individually.

There was also a concern that ethnic group demographics may change more rapidly than can be captured adequately by a 10-yearly census; for example, in relation to possible political change such as changes in policy on immigration and asylum, or exit from the EU. The relative healthcare need of the different ethnic groups may also change over time, due to (for example) education, programmes to reduce inequity, and generational cultural change. The subgroup recommends that ethnicity is re-examined in future reviews if better ways of capturing it can be identified.

#### 6.1.12 Did Not Attend (DNA) rate

A recent study [11] has shown that the outpatient ‘Did Not Attend’ (DNA) rate (as a percentage of total outpatient appointments) is highest among those living in more deprived and / or urban areas. It was also found that the patterning of DNA has been relatively stable for the past 10 years.

The DNA variable was included in the final list of potential indicators, and proved to be a strong indicator of utilisation; however, the subgroup decided to exclude it due to concerns that as a “supply-side” variable it is not appropriate and may be seen to set up a perverse incentive. Improving DNA rates is also a key performance indicator for boards and so improvement in DNA rates is intended, and this may affect the stability of the formula.

#### 6.1.13 Low birth weight births

Data was available from ISD on numbers of births with low birth weight. This was included because of its well-documented relationship with deprivation [12].

Table 7: List of candidate indicator variables entered into the selection process. Dementia prescribing, HRIs, DNA and Unpaid care were excluded at a later stage for reasons relating to TAGRA’s core criteria.

|  |  |
| --- | --- |
| **Variable** | **Details** |
| Low birth weight births | 3 financial years’ data from ISD Maternity team (10/11—12/13 – the three most recent years’ data available). Expressed as a fraction of average population over 3 years (MYEs 2011, 2012, 2013). |
| Death rate 0-74 all causes | 5 financial years’ NRS death records (09/10 – 13/14). Expressed as a fraction of average population over middle 3 years (MYEs – 2011, 2012, 2013) since 2010 MYE not available until Spring 2016.  Cause of death selected using ICD10 codes:   * Cancer C00--D48 * CHD I20--I25 * Stroke I61, I63, I64 |
| Death rate 0-74 Cancer |
| Death rate 0-74 CHD |
| Death rate 0-74 Stroke |
| All cause SMR 0-64 | Standardised mortality ratios with different causes of death.  Using 5 financial years’ GRO death records (09/10 – 13/14). SMR calculated using average population over middle 3 years (MYEs – 2011, 2012, 2013) since 2010 MYE not available until Spring 2016.  Cause of death selected using ICD10 codes:   * Cancer C00--D48 * Heart disease I00--I99 * Respiratory J00--J99 * Digestive K00--K93 * External Causes V\_\_--Y\_\_ * Other – any other codes |
| All cause SMR 0-69 |
| All cause SMR 0-74 |
| Cancer SMR 0-64 |
| Cancer SMR 0-69 |
| Cancer SMR 0-74 |
| Heart Disease SMR 0-64 |
| Heart Disease SMR 0-69 |
| Heart Disease SMR 0-74 |
| Respiratory SMR 0-64 |
| Respiratory SMR 0-69 |
| Respiratory SMR 0-74 |
| Digestive System SMR 0-64 |
| Digestive System SMR 0-69 |
| Digestive System SMR 0-74 |
| External Causes SMR 0-64 |
| External Causes SMR 0-69 |
| External Causes SMR 0-74 |
| Other SMR 0-64 |
| Other SMR 0-69 |
| Other SMR 0-74 |
| High Resource Individual counts | Count of individuals belonging to the group of ~100,000 highest resource-users that account for 50% of the total resource. Based on 3 financial years’ data from ISD IRF team (11/12 – 13/14). Expressed as a fraction of average population over 3 years (MYEs – 2012, 2013, 2014). |
| Did Not Attend counts – as a fraction of all OP appointments | 3 financial years’ data from ISD SC team (11/12 – 13/14). Expressed as a fraction of total outpatient appointments over same 3 financial years. |
| Did Not Attend counts – ratio to data zone population | 3 financial years’ data from ISD SC team (11/12 – 13/14). Average population over 3 years (MYEs – 2012, 2013, 2014). Included because of its link to deprivation. |
| Patients receiving Diabetes prescriptions | 3 financial years’ data from ISD Prescribing team (11/12 – 13/14). Expressed as a fraction of average population over 3 years (MYEs – 2012, 2013, 2014).  Dementia - includes all drugs in BNF section 4.11; Diabetes - includes all insulin and antidiabetic drugs; Respiratory - includes all lama, laba and high strength steroid inhalers |
| Patients receiving Dementia prescriptions |
| Patients receiving Respiratory prescriptions |
| Long-term illness | Data from Census 2011 question 20 – standardised by age and sex using 2011 MYE population. |
| Mental health condition | Data from Census 2011 question 20 – standardised by age and sex using 2011 MYE population. |
| Limiting long-term illness (LLTI) – limited a lot | Data from Census 2011 question 21 – standardised by age and sex using 2011 MYE population. Uses number of respondents answering ‘Yes – a lot’. |
| Limiting long-term illness (LLTI) – limited a little or a lot | Data from Census 2011 question 21 – standardised by age and sex using 2011 MYE population. Uses number of respondents answering ‘Yes’ (including both ‘a little’ and ‘a lot’). |
| Long-term sick and not seeking work | Data from Census 2011 questions 24-28 – standardised by age and sex using 2011 MYE population. |
| Older people living alone – 65 and over | Data from Census 2011 – standardised by age and sex using 2011 MYE population. |
| Older people living alone – 70 and over |
| Older people living alone – 75 and over |
| Older people living alone – 80 and over |
| Older people living alone – 85 and over |
| Older people living alone – 90 and over |
| Unpaid care provision – 1 hour or more | Data from Census 2011 question 9 – standardised by age and sex using 2011 MYE population. Uses number of respondents answering yes (with various numbers of hours). |
| Unpaid care provision – 20 hours or more |
| Unpaid care provision – 35 hours or more |
| Unpaid care provision – 50 hours or more |
| General health – very bad | Data from Census 2011 question 19 – standardised by age and sex using 2011 MYE population. Uses number of respondents answering ‘Very bad’. |
| General health – bad or very bad | Data from Census 2011 question 19 – standardised by age and sex using 2011 MYE population. Uses number of respondents answering ‘Bad’ or ‘Very bad’. |
| Education level – no qualifications | Data from Census 2011 question 23 – standardised by age and sex using 2011 MYE population. Level 1 refers to ‘O’ grades or similar; level 2 refers to Highers or similar (<http://www.scotlandscensus.gov.uk/variables-classification/highest-level-qualification>) |
| Education level – level 1 and below |
| Education level – level 2 and below |
| Job seekers | Data from Census 2011 question 25 – standardised by age and sex using 2011 MYE population (possibly economically active population only – try both). |
| Ethnicity I: all minority populations (including White minorities) | Data from Census 2011 question 15 – simple fraction of 2011 MYE population. The variables based on Gypsy/Traveller population, Pakistani population, and population of both groups were calculated at intermediate zone level to avoid large numbers of zero counts. |
| Ethnicity II: all populations with worse than average health (Gypsy/Traveller and Pakistani). |
| Ethnicity III: Gypsy/Traveller population |
| Ethnicity IV: Pakistani population |
| Ethnicity V: all populations with better than average health |

### 6.2 Indicator selection methodology

The Acute MLC review followed broadly the same indicator selection methodology as used in the original NRAC review. However, since the indicators of need were eventually to be combined into a single ‘needs index’ (as the sum of the Z-scores of the indicators) for use in the NRAC formula, it was agreed that the selection process should look for the best-performing *index* options, rather than constructing indices from the best-performing sets of *separate* variables (as had been done in 2007). These two approaches give different results.

The aim was to identify the ‘best’ single-variable index and the best two-, three- and four-variable index options for the subgroup to consider. Adjusted R2 was used as the metric to compare the index options’ performance.

The methodology is outlined in the sub-sections below. Figure 1 shows the overall methodology for the indicator selection process in the form of a flowchart.

#### 6.2.1 Stage one: elimination of near-duplicates

High numbers of data zones with zero counts for a potential explanatory variable are problematic for predictive modelling – the linear model does not fit well and its parameters are not well constrained. As a first step, the number of data zones with a ‘zero’ count was examined for all potential candidate variables, and variables with more than 30% of data zones having zero counts (2,093 or more out of 6,976 data zones) were eliminated.

Then, each candidate variable was allocated to an appropriate ‘topic’, so that unrelated variables (which are unlikely to be highly correlated) were held in separate topics. This reduced the number of inter-correlation calculations required. Four topics were used: (1) Births and deaths, (2) Health / Morbidity, (3) Unpaid care and older people living alone, and (4) Deprivation. As far as possible, this grouping follows that of the NRAC 2007 review (e.g. low birth weight births were included with the SMR and death rate variables). Unpaid care and older people living alone were grouped together because both care provision and an absence of care where needed (as may be the case where older people live alone) were anticipated to have a similar link to Acute healthcare need.

For each topic, the inter-correlations between the variables were computed. Subgroups were formed within each topic, each comprising highly inter-correlated variables (i.e. near duplicates). Variables that were not highly correlated with any other variables in their topic were automatically retained.

For each subgroup, the variables’ correlations with the cost ratios were computed; the variable that had the highest correlation with the cost ratios across most diagnostic groups was retained and the remaining ‘near duplicates’ removed.

#### 6.2.2 Stage two: identification of best diagnostic group-specific index options

The best-performing ‘specific’ index options (for each diagnostic group) were derived through exhaustive testing, to identify the best single-variable index and the best two-, three- and four-variable index options for the subgroup to consider. The steps were as follows:

* The best single-variable model is the variable with the highest adjusted R2 on its own for that diagnostic group, from the retained list.
* To identify the best two-variable model, all possible pairs of variables were combined into indices and tested in regressions, and the index with the highest adjusted R2 selected.
* The same exhaustive testing was carried out to find the best 3- and 4-variable models.

#### 6.2.3 Stage three: identification of best common index options

A ‘common’ needs index is currently used in the MLC adjustment for all Acute diagnostic groups, and having decided to retain this approach (see section 4.3), the best candidates for a common index were also identified.

The best ‘specific’ options for the different diagnostic groups were examined and the ‘common’ best 1, 2, 3 and 4-indicator models were simply taken to be those most commonly-occurring across the diagnostic groups. The best index for the ‘Whole Acute’ option was then compared with this, for validation.

To help the subgroup decide between the final set of options, adjusted R2 was examined again, and the models were also evaluated in predictive mode: predicted cost ratios were generated (in the same way as in the NRAC formula), which were then compared with a 1-year cost ratio based on 2014/15 data, using the residual sum of squares (RSS). The 2014/15 cost ratio represented the ‘future’ observation that each model was trying to predict.

**C:\Users\chrism29\Downloads\Index construction methodology (latest).png**

Figure 1: Flowchart showing final agreed methodology for choosing index options.

### 6.3 Results of the selection process

Results from the selection process were presented to the subgroup in papers TAMLC40, TAMLC43, TAMLC47, TAMLC49 and TAMLC53.

Several variables were removed due to having large numbers of zeros. Most of these were death rates or SMRs based on specific causes of death. Diabetes and Respiratory prescribing variables were also eliminated due to large numbers of zeros. Some of the Ethnicity variables had large numbers of zeros, but as detailed in section 6.1.11, the affected Ethnicity variables were recalculated at intermediate zone level to mitigate this.

After the near-duplicates analysis, the list of retained variables was as follows:

* All-cause SMR <75
* Cancer SMR <75
* Heart SMR <75
* Other SMR <70
* Limiting long-term illness (LLTI) – limited a little or a lot
* Living alone ≥70
* Living alone ≥90
* Unpaid care ≥ 20 hours
* Education – level 2 and below
* DNA counts – fraction of appointments
* Low birth weight births
* Patients receiving Dementia prescriptions
* High Resource Individuals (HRIs)
* Long-term sick and not seeking work
* Ethnicity V: Ethnic populations with better than average health
* Ethnicity IV: Pakistani populations
* Ethnicity III: Gypsy/traveller population

At this point the Dementia prescribing and HRI variables were also excluded, for the reasons given in sections 6.1.9 and 6.1.10, respectively.

Table 8 shows the results of the exhaustive testing to identify the best diagnostic group-specific index options, for 1, 2, 3 or 4 indicators.

Table 8: Best-performing index options using 1, 2, 3 and 4 variables

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Whole Acute top 4** | **Cancer specific top 4** | **Heart specific top 4** |
| **Top 1**  **Top 2**  **Top 3**  **Top 4** | LLTI  LLTI, All-cause SMR <75  LLTI, All-cause SMR <75, Unpaid care  LLTI, All-cause SMR <75, Unpaid care, DNA | Cancer SMR <75  Cancer SMR <75, All-cause SMR <75  Cancer SMR <75, All-cause SMR <75, Unpaid care  Cancer SMR <75, All-cause SMR <75, Unpaid care, Living alone 90 | LLTI  LLTI, DNA  LLTI, DNA, All-cause SMR <75  LLTI, All-cause SMR <75, Unpaid care, DNA |
|  | **Digestive specific top 4** | **Injury specific top 4** | **Other specific top 4** |
| **Top 1**  **Top 2**  **Top 3**  **Top 4** | LLTI  LLTI, All-cause SMR <75  LLTI, All-cause SMR <75, Unpaid care  LLTI, All-cause SMR <75, Unpaid care, DNA | LLTI  LLTI, All-cause SMR <75  LLTI, All-cause SMR <75, DNA  LLTI, All-cause SMR <75, Unpaid care, DNA | LLTI  LLTI, All-cause SMR <75  LLTI, All-cause SMR <75, Unpaid care  LLTI, All-cause SMR <75, Unpaid care, DNA |
|  | **Respiratory specific top 4** | **Outpatients specific top 4** |  |
| **Top 1**  **Top 2**  **Top 3**  **Top 4** | LLTI  LLTI, All-cause SMR <75  LLTI, All-cause SMR <75, Unpaid care  LLTI, All-cause SMR <75, Unpaid care, DNA | LLTI  LLTI, Unpaid care  LLTI, Unpaid care, Education level  LLTI, Unpaid care, Education level, DNA |

The best ‘common’ index options comprising 1, 2, 3 and 4 variables were taken to be:

* LLTI (best single-variable index for 6 out of 7 diagnostic groups and for Whole Acute)
* LLTI, All-cause SMR <75 (best 2-variable index for 4 out of 7 diagnostic groups and for Whole Acute)
* LLTI, All-cause SMR <75, Unpaid care (best 3-variable index for 3 out of 7 diagnostic groups and for Whole Acute)
* LLTI, All-cause SMR <75, Unpaid care, DNA (best 4-variable index for 5 out of 7 diagnostic groups and for Whole Acute)

Many of the best diagnostic group-specific options were identical or contained at least some common variables, and most are subsets of the same four variables: LLTI, All-cause SMR <75, Unpaid care, and DNA.

Where the best specific index differed from the common one, the predictive power did not seem to be substantially higher using the specific index. This was true even for Cancer, which had consistently higher adjusted R2 for its specific index options, all of which contained Cancer SMR <75 – a variable with obvious relevance and face validity for Cancer. Perhaps surprisingly, the specific index options for Cancer actually performed slightly *less* well than the common indices in predictive mode (although the difference was quite small).

Out of all the variables, Cancer SMR <75 is most clearly related to the dependent variable: due to temporally overlapping data, it is very likely that some of the cancer deaths counted in this variable were associated with Acute healthcare costs prior to death that are reflected in the cost ratios used to fit the model. This could explain why there appears to be a much stronger relationship between the indicator and the cost ratios used in fitting the model than between the indicator and *future* costs.

There was no real evidence that different condition groups have different needs drivers (at least, from the set of candidate variables analysed in this review), nor that diagnostic group-specific indices will perform better over the long term and justify the extra complexity. The subgroup therefore decided that the Acute MLC adjustment should continue using a common index across all diagnostic groups.

At this point, DNA and Unpaid care were also excluded, for reasons relating to the TAGRA core criteria, as discussed in sections 6.1.12 and 6.1.7 respectively. Ethnicity variables were some of the best-performing indicators remaining, and were favoured by the subgroup *a priori* on equity grounds; however, the subgroup also decided to exclude Ethnicity V, for the reasons discussed in section 6.1.11. Because of this latter removal, Ethnicity II, III and IV were reintroduced and included in the final set of options instead. These options were:

* LLTI, All-cause SMR <75
* LLTI, All-cause SMR <75, Ethnicity II (groups with poorer than average health)
* LLTI, All-cause SMR <75, Ethnicity III (Gypsy/Traveller populations)
* LLTI, All-cause SMR <75, Ethnicity IV (Pakistani populations)

Table 9 and Table 10 show that the addition of ethnicity variables to the two-variable model – LLTI + All-cause SMR <75 – appears to result in poorer performance, in terms of both explanatory power (adjusted R2) and predictive power (RSS). While adding variables in a multiple regression would always increase the adjusted R2, even if only marginally, this is not the case when adding variables to a combined index which is then used as a single variable.

This poorer statistical performance of index options including ethnicity, combined with the other concerns about the ethnicity variables detailed in section 6.1.11, led the subgroup to conclude that the best option for the Acute MLC index remains the two-variable option selected in the 2007 NRAC review: LLTI and All-cause SMR <75.

Recommendation 8: retain Limiting Long-Term Illness and All-cause Standardised Mortality Ratio ages 0-74 as the needs indicators.

Table 9: Adjusted R2 for various index options, by diagnostic group. The average, weighted by diagnostic group spend, is shown in the final column.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | **Cancer** | **Heart** | **Digestive** | **Injury** | **Other** | **Respiratory** | **Outpatients** | **Weighted**  **average** |
| LLTI, All-cause SMR <75 | 10.9% | 21.0% | 38.6% | 26.0% | 45.1% | 38.2% | 49.3% | 36.9% |
| LLTI, All-cause SMR <75, Ethnicity IV | 10.0% | 17.7% | 34.1% | 22.8% | 39.4% | 33.4% | 46.4% | 32.9% |
| LLTI, All-cause SMR <75, Ethnicity III | 9.9% | 17.5% | 33.3% | 21.2% | 39.1% | 33.2% | 46.9% | 32.6% |
| LLTI, All-cause SMR <75, Ethnicity II | 9.9% | 17.5% | 33.3% | 21.0% | 39.1% | 33.1% | 47.0% | 32.6% |

Table 10: RSS obtained from comparing predictions derived from the index options with the 2014/15 cost ratios. Lower values indicate the predictions are closer to the observed value.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | **Cancer** | **Heart** | **Digestive** | **Injury** | **Other** | **Respiratory** | **Outpatients** |
| LLTI, All-cause SMR <75 | 3264 | 3618 | 2200 | 2852 | 828 | 3569 | 396 |
| LLTI, All-cause SMR <75, Ethnicity IV | 3292 | 3632 | 2333 | 2911 | 913 | 3793 | 407 |
| LLTI, All-cause SMR <75, Ethnicity III | 3272 | 3657 | 2283 | 2936 | 882 | 3648 | 396 |
| LLTI, All-cause SMR <75, Ethnicity II | 3272 | 3661 | 2279 | 2937 | 881 | 3646 | 396 |

## CHAPTER 7: ADDITIONAL VARIABLES

### 7.1 Prison indicator

A number of outliers were apparent in scatter plots, for which the cost ratios were substantially larger than would be expected from the linear trend. This was particularly the case for Outpatients. There were no influential points, however; i.e. no outliers that altered the slope of the fitted line substantially by their inclusion.

Investigation into the two largest cost ratio outliers for Outpatients revealed that the data zones both contained prisons. For both data zones, males between the ages of 20 and 49 accounted for the largest part of expenditure, which suggested the prisoner Outpatient activity was indeed the reason for these extreme values.

The subgroup considered whether the few data zones containing prisons should be removed from the regression analysis. Despite the fact that they are not influential points, they represent concentrated populations with healthcare needs that are likely to have a very different relationship with the indicators of need as compared to the general population.

However, the funding of Acute healthcare for prisoners is covered by the NRAC formula (primary care within prisons is covered by a separate allocation process). Another suggestion was to include a ‘dummy’ variable (i.e. a binary indicator) to represent the presence of a prison within a data zone, so that the additional cost associated with prisoner Acute care could be modelled (simply, as a constant amount to be added to a data zone containing a prison) and included in the target shares. A prison dummy variable was added to the reference model to test this idea; its coefficient turned out only to be significantly different from zero for Outpatients. This meant that, according to the data, prisons are not associated with higher utilisation for Acute Inpatient and Daycase services – Acute Outpatient services alone are utilised at a higher rate.

The subgroup agreed to implement the prison indicator for Outpatients. This is to be entered into the regression separately from the combined needs index, in the same way as the urban-rural indicators have been entered for Maternity. Its regression coefficient will then be used in the prediction of need.

Recommendation 9: introduce a prison indicator (a binary variable) for Acute Outpatients, to allow additional need associated with prisons to be allocated for.

### 7.2 Urban-rural category indicators

It was decided during the NRAC review in 2007 that urban-rural (UR) indicators should be included in the MLC model for Maternity. Although the Excess Costs component of the formula accounts for UR effects, it only does so through allowing for a *unit cost* that varies by UR setting – it does not account for differing activity levels (such as lengths of stay) between regions. Including UR markers in the MLC model accomplishes this, where it is found to be the case.

After the new Acute model was decided upon, possible urban-rural effects were investigated; following the method of previous reviews, this was done by including UR markers as extra binary variables in the regression of cost ratios upon the needs index and supply model, which effectively adjust the healthcare need of each rural area within the same urban-rural category by a constant amount. The analysis involved examining the effect on the coefficient of the needs index, the adjusted R2, and the predictive power (measured by RSS).

Following the previous review methods, four different UR classifications have been tested: a 2-fold, a 4-fold, a 6-fold and an 8-fold classification. Definitions of the classifications are given in Annex B.

The inclusion of urban-rural markers (of any classification) did not produce significant differences in the coefficient of the proposed Acute needs index at the 5% level. The adjusted R2 was increased only very slightly when urban-rural markers are included. The predictive power also did not improve substantially when including the urban-rural markers.

Many coefficients of the urban-rural indicators were not significant, particularly for Cancer, Injury and Outpatients. There was also no clear pattern. If activity levels varied significantly with urban-rural setting, then it might be expected, for example, that the sign of the coefficient (for a given diagnostic group) was consistently negative for urban categories and positive for more rural categories. No such pattern was evident.

In conclusion, the proposed Acute MLC model appears to perform similarly well across all urban-rural settings. The subgroup members were in agreement that no significant improvement could be anticipated from including urban-rural indicators in the model.

Recommendation 10: do not include urban-rural category indicators in the Acute MLC model.

## CHAPTER 8: INVESTIGATION OF UNMET NEED

The NRAC formula relies on health service activity data as a basic proxy for the need for healthcare services. It is therefore important to check for evidence of any socio-economic inequities in healthcare utilisation and, where appropriate, to adjust the formula to reflect such unmet need. There is currently an unmet need adjustment for Heart based on excluding the 25% most deprived intermediate zones, based on the SIMD income domain, from the regression.

Prior to the indicator selection work, preliminary testing for unmet need was carried out in four diagnostic groups – Cancer, Digestive, Heart and Respiratory – based on the reference model at the 2001 data zones. This was done to explore the utility of different methodologies and agree on the methodology to be used in the eventual analysis with the new MLC model.

Several methodologies were considered:

* Simple shortfall method, used in the Arbuthnott review [13]. This adds a ‘spline’ term to the regression to allow for a different slope in the regression line at the high or low end (or both) of the needs index. This can be used to test for under-utilisation, based on the assumption that there would be a constant linear relationship between the cost ratio and the needs index if there was no unmet need.
* 2007 shortfall method, used in the 2007 NRAC review [4]. It allows a subset of the data zones to have a different slope and/or intercept, and the subset can be chosen based on any characteristic such as high SIMD ranking.
* Two-step shortfall method [13]. This method involves a two-stage process using independent morbidity data to define expected levels of utilisation. Scottish Health e-Survey (SHeS) data was trialled for this purpose, using surveys from 2008 to 2011 for which there were around 37,000 respondents. The data were coded by ICD-10 which allowed a match to diagnostic groups.

Following the initial analysis, the subgroup decided that the two-step shortfall method should not be pursued further: the SHeS data was very sparse and so the modelling between the Acute index and SHeS morbidity was not well constrained; there were also concerns about the face validity of its non-linear model given that a linear model had been chosen for the MLC adjustment.

The subgroup instead agreed to look for any unmet need at the high end of the Acute needs index using the simple shortfall method, and to use the 2007 shortfall method to look for unmet need based on certain other characteristics: deprivation, urban-rural setting and ethnicity. This approach allowed for the possibility of finding unmet need along several different ‘dimensions’ where it may plausibly exist.

### 8.1 Methodology

The aim of shortfall methods is to help identify a subset of data zones in which the utilisation is lower than it should be, due to unmet need. If these affected data zones influence the slope of the fitted linear model, they can be excluded from the regression in the eventual implementation of the MLC model. The resulting ‘adjusted’ slope is then used to predict need for *all* data zones.

The simple shortfall method looks for evidence of unmet need at the high end of the needs index. This is done by fitting a model with a break of slope, as illustrated in Figure 2 (top), to identify whether the effective gradient of the relationship between needs index and cost decreases significantly at the high end of the needs index.

The 2007 shortfall method (Figure 2, bottom) looks for systematically lower utilisation in the subset of data zones with the characteristic of interest (high deprivation, the most urban / rural settings, or high proportions of particular ethnic group populations) compared to the bulk of data zones. This is achieved by allowing both the slope and intercept to be different in the selected data zones, effectively fitting two linear models (L1 and L2).

Arguably, the 2007 shortfall method has more face validity than the simple shortfall method, as it does not rely so heavily on the assumption of a linear relationship between needs index and cost. Evidence of non-linearity from the simple shortfall method is therefore not sufficient to infer unmet need, and should be accompanied by further reasoning (i.e. face validity) if an unmet need adjustment is to be considered.

#### 8.1.1 Simple shortfall method

In the simple shortfall method, a spline variable is added to the linear model to allow the slope to change at a cut-off point. The spline variable is defined using the Acute needs index values (*xi*, where *i* refers to the data zone), as:

*xiH* = *xi – k* if *xi > k*

=0 if *xi ≤ k*

where *k* is the needs index value corresponding to the cut-off point being used.

First of all, the cut-off point must be set: all possible cut-off points between 5% and 30% are trialled in the above model and the one that yields the highest explanatory power, i.e. the highest adjusted R2 value, is adopted for the analysis.

Then, if the spline term’s coefficient is significant and negative, this indicates that the slope decreases at the high end of the needs index, i.e. there is evidence of unmet need based on the linearity assumption. Furthermore, if the slope *β1* in Figure 2 (top) is significantly higher than the ‘unadjusted’ slope of the simple linear model fitted to all data zones, this shows that the overall fitted model is affected by the potential unmet need; and therefore, that an unmet need adjustment – to remove the affected data zones from the regression – would be effective.

**Utilisation**

cut-off point

slope *β2*

slope *β1*

**Acute needs index**

L2

**Utilisation**

**Acute needs index**

L1

subset of data zones with high deprivation (or other characteristic)

Figure 2: Illustration of the simple shortfall method (top) vs the 2007 shortfall method (bottom).

#### 8.1.2 2007 shortfall method

In the 2007 shortfall method, a binary variable indicates the data zones with the characteristic of interest. Two terms – the binary variable itself (*binaryi*) and an ‘interaction’ term (*needs indexi* × *binaryi*) – are added to the linear model:

For those data zones that have ‘binary’ equal to 0, i.e. the majority, the linear model has intercept *β0* and slope *β16*; but for the data zones with ‘binary’ equal to 1, i.e. the areas with the characteristic of interest, the linear model is allowed to be different: the intercept is (*β0* + *β17*), and the slope is (*β16* + *β18*). A range of thresholds for the characteristic of interest are tested and the one yielding the highest R2 is adopted, as for the simple shortfall method.

If either of the additional terms is significant, there is evidence of different utilisation patterns in the areas with the characteristic of interest – significance in the binary variable indicates a different intercept, i.e. a constant shift in utilisation up or down from the level expected based on the needs index; and a significant interaction term indicates a different slope. We may infer an unmet need adjustment is needed if (1) the binary variable is significant, with a negative coefficient, implying lower utilisation in the data zones with the characteristic of interest; *and* (2) the needs index coefficient is significantly changed by this – in other words, if the slope of L1 in Figure 2 (bottom) is significantly different from the ‘unadjusted’ slope of the simple linear model fitted to all data zones.

### 8.2 Results

#### 8.2.1 Unmet need related to the needs index

There was found to be a significant shortfall from the expected utilisation in the data zones with the highest needs index values, for Heart (with cut-off point 30% selected by the highest-R2 criterion) and Other (with cut-off point 5% selected), which significantly altered the slope of the fitted line as compared to when these data zones were excluded from the regression.

From these results, a decision was made to recommend an unmet need adjustment for the Heart diagnostic group, based on excluding the 30% of data zones with the highest needs index values from the regression. This was agreed because:

* An unmet need adjustment for Heart has face validity: untreated heart disease can cause sudden death (with potentially no prior contact with the health services), and so it is plausible that lower levels of resource are utilised where there is unmet need for diagnosis and treatment. This may not be true for conditions such as Cancer, where unmet need might result in late interventions that would still be costly (and perhaps even more costly than early treatment) and so would perhaps not be expected to show up as under-utilisation.
* Supplementary analysis showed that the standardised mortality ratio for heart-related deaths did not tail off at the upper end of the needs index in a similar way to cost, but actually showed an increase in gradient, which further suggests need that is not being met.
* There is currently an unmet need adjustment for Heart, which suggests that the above observations are consistent and robust. However, this does also highlight that other mechanisms are needed to address the unmet need, and this should be noted by TAGRA.

Similarly, the subgroup decided to recommend an unmet need adjustment for the Other diagnostic group, based on excluding the 5% of data zones with the highest needs index values from the regression. It was more difficult to judge whether the non-linearity should be interpreted as unmet need in the case of Other, given the varied mixture of condition groups comprised within Other (Table 11). However, an apparent shortfall was observed over the whole range of cut-off points examined, indicating that the effect is robust. Moreover, a cut-off of 5% represents a cautious approach.

Recommendation 11: implement an unmet need adjustment for Heart based on excluding the 30% of data zones with the highest needs index values from the regression, and an unmet need adjustment for Other based on excluding the 5% of data zones with the highest needs index values.

Table 11: ICD-10 code groups within Acute Other diagnostic group.

|  |  |  |
| --- | --- | --- |
| **ICD-10 group** | **Actual spend 2013/14** | **Number of episodes 2013/14** |
| Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified | £319m | 204,912 |
| Diseases of the musculoskeletal system and connective tissue | £221m | 92,259 |
| Diseases of the genito-urinary system | £198m | 101,434 |
| Diseases of the nervous system | £92m | 38,272 |
| Certain infectious and parasitic diseases | £87m | 40,811 |
| Factors influencing health status and contact with health services | £81m | 57,698 |
| Diseases of the skin and subcutaneous tissue | £71m | 32,433 |
| *missing* | £70m | 25,176 |
| Diseases of the eye and adnexa | £70m | 50,347 |
| Endocrine, nutritional and metabolic diseases | £51m | 28,893 |
| Mental and behavioural disorders | £38m | 16,406 |
| Congenital malformations, deformations and chromosomal abnormalities | £30m | 9,669 |
| Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism | £28m | 21,625 |
| Diseases of the ear and mastoid process | £9m | 6,653 |
| Certain conditions originating in the perinatal period | £2m | 1,281 |

#### 8.2.2 Unmet need related to ethnicity

For the analysis, the total Pakistani & Gypsy/Traveller population as a percentage of the total for each data zone, from the 2011 census, was used as a measure of ethnicity. These ethnic groups were chosen, as in prior ethnicity analysis, because they were independently identified as having worse-than-average health outcomes [10]. The data zones were then ranked based on their percentage of Pakistani & Gypsy/Traveller population. Areas were categorised based on a number of cut-points: successively, the 1%, 5%, 10%, 15%, 20%, and 25% of the total number of data zones with the highest percentage of Pakistani & Gypsy/Traveller population.

Although the shortfall model’s extra terms were sometimes significant, the needs index coefficient (i.e. the slope of the overall fitted model) was not significantly changed in the shortfall model for any diagnostic group. An unmet need adjustment relating to ethnicity would therefore not be needed.

#### 8.2.3 Unmet need related to urban/rural setting

Remoteness and rurality was measured using the 6-fold Scottish Government Urban Rural Classification (see Annex B for details of this). As per the 2007 NRAC analysis, this was converted into a 3-fold classification by grouping categories as follows: categories 1 & 2 (Urban), category 6 (Remote and Rural) and between them the remaining categories 3, 4 & 5 (Other). “Urban” contains 4,837 data zones (69%), “Remote and Rural” contains 425 (6%), and the remainder is contained within the “Other” category.

Two distinct comparisons were carried out. Firstly, remote and rural areas were compared to all other areas, and then all non-urban categories were compared to urban areas. This meant that unmet need was tested for at both ends of the urban-rural spectrum.

The needs index coefficient was not substantially changed when including the additional terms – the confidence intervals overlapped. The smallest degree of overlap was for Injury, using the shortfall model that separates out urban areas. However, the binary variable was not significant in that case, indicating that there was no under- or over-utilisation as compared to the rest of the country. This was not regarded as unmet need.

#### 8.2.4 Unmet need related to deprivation (SIMD)

It was not possible to look for evidence of unmet need using SIMD, as it is currently unavailable at the 2011 data zones at which the analysis for the review was carried out. SIMD 2016 will be provided at the 2011 data zones and is expected to be released in August 2016.

## CHAPTER 9: HEALTH INEQUALITIES IMPACT ASSESSMENT

The Acute MLC subgroup was keen to ensure that a Health Inequalities Impact Assessment (HIIA) was undertaken during its work. Recognising good practice indicates that HIIA should be considered from the outset. A presentation about HIIA was given at an early meeting of the committee and a representative from NHS Health Scotland was invited to join the subgroup.

HIIA is an integrated approach to impact assessment that brings together consideration of equality, social determinants of health and human rights to achieve effective policy making and service development. The purpose of HIIA is to provide a systematic method to identify and reduce barriers to equitable implementation of the service or policy at a stage in planning where changes can be made. The subgroup aimed to embed equity considerations throughout the Acute MLC review and adapted NHS Health Scotland’s HIIA materials [14] to devise a tailored process to support systematic consideration of equity throughout. This comprised three components: regular, informal assessment of proposals and analyses at subgroup meetings; a focussed HIIA workshop discussion at the midpoint; and a process of finalising results following further analyses.

### 9.1 Consideration of equity throughout the review

The first component – of ongoing consideration of equity within regular meetings – led to an amendment of the existing TAGRA core criteria, to expand the definition of equity to specifically mention variation in need across population groups. Later discussions helped to identify additional inequality indicators for potential inclusion including Child Poverty, Child Pedestrian Casualties, numbers of asylum seekers and refugees, Roma community, Did Not Attend (DNA) rates, life expectancy, and the gender gap in life expectancy. Some of these were rejected due to the potential for double counting, having a low impact on Acute MLC, or that data were unavailable. These explorations concluded that DNA rates and life expectancy should be included in the list of potential new indicators and that data zone was the preferred geographical level over intermediate zones for reasons of responsiveness to deprivation, equality and transparency.

### 9.2 HIIA workshop

The second component of the HIIA was to hold a workshop discussion at an interim stage approximately half-way through the review to offer the group an opportunity for systematic exploration of the potential impacts of the formula on equality, inequality and human rights. Additional equality experts were invited to contribute to the discussion to review progress to date and to help the subgroup explore further questions relating to equity at a point where any omissions or negative impact could be remedied. The subgroup acknowledged limitations of the formula’s potential for strengthening equity including, first, that the formula’s impact is mediated through individual NHS Board allocation processes; second, that there are likely to be gaps in routinely collected data for understanding diversity of health needs in small population groups; and finally, that actions to further explore or mitigate some potential harms identified might lie outwith the scope of the subgroup or the allocation formula.

The workshop discussion ruled out the legally protected characteristics of religion and belief and marriage and civil partnership as useful indicators for need, and considered that suitable data on need for service use was unlikely to be available for sexual orientation, literacy, or looked after and accommodated children and young people. Issues relating to staff were covered elsewhere by the formula (the Unavoidable Excess Costs index, which takes account of the excess costs of supplying health services in different urban-rural areas). Gender, age, deprivation and rurality were already under review by the subgroup. People over 75 living alone and the prevalence of long-term limiting conditions were suggested as potential indicators (long term limiting conditions were already included). Carers were identified as a group already noted in health policy and of interest as a potential indicator for additional healthcare need, for two reasons. Firstly, being a carer can impact negatively on health, and secondly, for those needing care, absence of a carer could have an impact on acute services for example in exacerbating delayed discharges. A discussion on human rights concluded that these were important but difficult to demonstrate specifically within the formula itself other than through strengthening equity as an explicit principle underpinning all aspects of the formula review processes.

Extremely marginalised groups whose specific needs are unlikely to be reflected in routine data were discussed. Groups identified included refugees and asylum seekers, people experiencing addictions and substance misuse, those who are homeless, those who are involved in the criminal justice system, and Gypsy/Travellers. Homelessness was explored in depth to identify data sources and impact of complexity of need in such groups on acute services. The discussion concluded that there may be a need for other ways to think about health care provision for such small groups for example through the High Resource Individuals initiative or through the Scottish Allocation Formula (SAF) for GP practices. Some of these small groups were included in discussions about ethnicity which was already being considered as an indicator but some questions were raised about this as outlined below.

The proposal to group together all ‘non-white’ ethnic groups was considered too crude as ‘non-white’ would fail to capture some groups such as Gypsy/Travellers and migrant workers who have some of the poorest health outcomes. Some ‘non-white’ ethnic groups have better health outcomes than average, others have worse. This could be for reasons not always necessarily directly related to ethnicity, for example, poor health could be related to poverty or a combination of factors rather than through identifying as belonging to a minority ethnic group. The ethnic profile of Scotland is changing quickly and the previous NRAC report shows that areas with higher proportions of ethnic minorities also have higher values of the main indicators of need. However, it is important to understand the context of data around ethnicity. For example, new immigrant populations often settle in the most disadvantaged areas but as they become more settled or by the second or third generation, they may no longer be concentrated in poorer areas and consequently health correlations change.

It was argued that ethnicity should be included as an indicator in some form because there is academic evidence of inequalities linked to ethnicity; however, the assumptions being used for ethnicity needed to be tested. The potential variables and the process of testing are described in Chapter 6.

Disability was discussed at the workshop. This raised the complexity of measurement, for example some people with life-long limiting conditions or co-morbidities may be high resource users but also may be effectively engaged in self-management. In addition, many conditions are age-related and there would be a need to distinguish between older people and disabled people who also happen to be older. It may be hard to disaggregate a number of cumulative factors all impacting on health such as disability, old age, living alone and living in a disadvantaged area.

The outcome of the HIIA workshop discussion was to take forward further analyses of ethnicity and to consider candidate variables for carers, adding to inclusion of long term limiting illness by exploring unpaid care.

### 9.3 HIIA findings

Four potential variables were identified through the extended HIIA process. DNA rates and life expectancy were proposed during the first component of the HIIA and DNA rates was added to the initial list of potential candidate variables for exploration. At a later stage the HIIA workshop identified unpaid care and ethnicity as candidates for inclusion. The third component of the HIIA was for the group to reach consensus on inclusion of the variables following analyses by the review team and agreement of the equality advisers, described below.

***Life expectancy:*** Life expectancy was proposed as a potential candidate inequality variable at an early stage in the review as it is commonly used in planning by NHS Boards, local authorities and community planning partnerships. Life expectancy information was available through the Scottish Public Health Observatory (ScotPHO) at intermediate zone level only, and for five time periods: 1994-98, 1996-2000, 1999-2003, 2001-05 and 2003-07. Standardised mortality ratios on the other hand can be calculated for any geography and time period. Life expectancy was therefore not included in the list of potential candidate variables.

***DNA rates:*** DNA rates were thought to be a strong indicator for utilisation and remained part of the candidate list until a late stage in the analysis. Concerns were raised that the DNA rate was a supply variable and that inclusion for cost allocation could incentivise a lack of action while at the same time, NHS Boards are expected to improve their DNA rates through performance measures. The group agreed that inclusion could potentially be counterproductive to strengthening equity.

***Unpaid care:*** The issue of carers was raised at the HIIA workshop. Carers were identified in policy as a population group with high levels of need, and were of interest to the group for the potential for unpaid care to mask some of the demand for acute services. However, if carers themselves developed a need for acute care, their charges might also need additional care. The absence of a dose-response relationship between unpaid care provision and higher acute costs in the data (described in section 6.1.7) was of concern as it was counterintuitive to the evidence of greater healthcare need in carers. The subgroup agreed that unpaid care should be re-examined in later reviews if more responsive ways of capturing need can be developed.

***Ethnicity:*** the search for an indicator for ethnicity proved to be complex and is described in Chapter 6 and also in Chapter 8 where ethnicity was included in the unmet needs analysis. Equality experts at the HIIA workshop advised against using ‘white’ and ‘non-white’ categories as these would not capture complex patterns of need across diverse ethnic groups. In addition, the subgroup was concerned that diversity in the Scottish population is rising and likely to change over the census periods. Extensive analyses failed to find an indicator that captured additional need related to ethnicity that would remain stable over time and / or that would impact to the level that would require adjustment to the formula. The subgroup recognised the theoretical link between ethnicity and acute healthcare need and the face validity of including an ethnicity variable but agreed that ethnicity could not be captured adequately in the formula. The subgroup recommended that ethnicity is re-examined in future reviews if more responsive ways of capturing need in relation to ethnicity for the population can be identified.

### 9.4 Conclusion

The subgroup used an adapted HIIA process in order to consider equality, inequality and human rights throughout the review process. Equality advisers participated from the early stages of the review to the end, arguably enabling the subgroup to be more involved in assessing inequality impact of their own work than would be possible in the more traditional process of a single workshop. The subgroup ran a focused HIIA workshop discussion at a point in the process where consideration of equality, inequality and human rights was already embedded in the process but where new issues identified could still be addressed within the timescale of the review. While none of the four inequality-related candidate variables were added to the final formula, the process modelled a commitment to analysing the impact of the formula on equity in acute healthcare. The benefits of considering equity throughout the process were that there were multiple opportunities to consider equality, inequality and human rights throughout and that the whole subgroup and review team were involved in all aspects of assessing the equity impact of the review at one or more points in the process. Limitations were identified in the responsiveness of routine population data to diversity of need in the population and the review recommended that two of the candidate variables be re-visited in future reviews if better ways of capturing need could be developed.

## CHAPTER 10: FURTHER RECOMMENDATIONS

### 10.1 Limitations of the present review and recommendations for the future

It is worth noting that in the recent review of the SAF formula, a single regression model was developed to account for both age-sex effects and additional needs. This differs from the NRAC formula structure, in which age and sex effects are accounted for separately and then the dependent variable in the additional needs model is age-sex standardised. This means that no interactions between age-sex and MLC factors are accounted for in NRAC, except (in the Mental Health & Learning Difficulties model) by the separation of age groups into different models. The Acute MLC review found some evidence of interactions (section 4.4) but concluded that an age split would not improve predictive power. It would have been outside the remit of the Acute MLC review to alter the overall structure of the NRAC formula, but TAGRA may wish to consider the merits of a single age-sex-MLC model in future developments.

Apart from the elimination of ‘near duplicates’ (section 6.2.1), inter-correlations between indicators were not examined as part of the selection process, and are not very important given the procedure within the NRAC formula of regressing cost ratios upon a combined index (as the sum of the z-scores of the needs indicators). Again, it was outside the remit of the present review to alter this practice, but it may be worth reviewing in future resource allocation developments, in which case indicator selection methodologies should be reviewed accordingly. In general, the 2007 process of indicator selection was refined considerably in this review (section 6.2) but could be further improved.

At the time the costing method was reviewed, PLICS was still in development, but there was general agreement that a move to full use of PLICS costs within the NRAC formula would be favoured in the future. More recently, PLICS has been used in the production of the Atkinson output submitted to the ONS for use in measurement of economic growth. The subgroup would therefore recommend that the use of full PLICS costs for Acute inpatients and daycases is considered by TAGRA in the near future.

The subgroup recognised the potential importance of multimorbidity and mental health issues as predictors of need for Acute services [7], but it was not possible to take full advantage of these insights due to a lack of detailed data on disease prevalence. The selected indicators, the limiting long-term illness ratio and the all-cause standardised mortality ratio (<75), essentially represent aggregated disease prevalence. Better data are needed to capture disease prevalence in more detail at data zone level in order to be able to assess these issues. In particular, primary care data from the Scottish Primary Care Information Resource (SPIRE) should be explored when it becomes available; prescribing data could also be further explored and linked to other data sets to provide information on additional co-morbidities. Improvement in the recording of co-morbidities in SMR01 should also be utilised.

The subgroup recommends that both ethnicity and unpaid care are re-examined as potential needs indicators in future reviews if better ways of capturing the need responsively can be identified.

## REFERENCES

[1] Delivering Fair Shares for Health in Scotland: The report of the NHSScotland Resource Allocation Committee, September 2007, available from <http://www.tagra.scot.nhs.uk/research>

[2] TAGRA Morbidity & Life Circumstances Subgroup, "NRAC formula - MLC adjustment: Mental Health and Learning Difficulties care programme: Final Report", December 2012, available from <http://www.tagra.scot.nhs.uk/subgroups/morbidity-and-life-circumstances/>

[3] Technical Report D: Review of the Resource Allocation Adjustment for Healthcare Needs Due to Morbidity and Life Circumstances and Other Factors, May 2006, available from <http://www.tagra.scot.nhs.uk/research>

[4] Addendum to Technical Report D: Review of the Resource Allocation Adjustment for Healthcare Needs Due to Morbidity and Life Circumstances and Other Factors: Additional Analyses Following Consultation, August 2007, available from <http://www.tagra.scot.nhs.uk/research>

[5] International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10), available from <http://apps.who.int/classifications/icd10/browse/2015/en>

[6] Michael Batty, “Urban Modelling: Algorithms, Calibrations, Predictions”, Cambridge University Press 1976. Available online at <http://www.casa.ucl.ac.uk/urbanmodelling/UrbanModelling.pdf> - the relevant section is ‘The Allocation of Urban Activities’, p.35 onwards.

[7] Karen Barnett et al., “Epidemiology of multimorbidity and implications for health care, research and medical education: a cross-sectional study”, The Lancet 380 (9836), p.37-43, July 2012.

[8] Scottish Government report “Scotland’s Carers”, March 2015, available from <http://www.gov.scot/Resource/0047/00473691.pdf>

[9] Scottish Government Social Research report “Caring in Scotland: Analysis of Existing Data Sources on Unpaid Carers in Scotland”, 2010, available from <http://www.gov.scot/Resource/Doc/319575/0102110.pdf>

[10] Scottish Government ScotStat report “Which ethnic groups have the poorest health? An Analysis of Health Inequality and Ethnicity in Scotland”, August 2015, available from <http://www.gov.scot/Publications/2015/08/7995/downloads>

[11] NHS Health Scotland report “Who is least likely to attend? An analysis of outpatient appointment ‘Did Not Attend’ (DNA) data in Scotland”, March 2015, available from <http://www.healthscotland.com/documents/25015.aspx>.

[12] Scottish Government Statistical Bulletin (Health & Social Care series), “Long-term Monitoring of Health Inequalities”, October 2014, available from <http://www.gov.scot/Resource/0046/00461784.pdf>

[13] McConnachie and Sutton, ‘Derivation of an Adjustment to the Arbuthnott Formula for Socioeconomic Inequities in Health Care’, 2004, available from <http://www.gov.scot/Publications/2004/02/18945/33395>.

[14] Health Inequalities Impact Assessment, NHS Health Scotland, available at <http://www.healthscotland.com/equalities/hiia/index.aspx>

## ANNEX A: ACUTE MLC SUBGROUP MEMBERSHIP

Karen Facey (Chair)

Diane Skåtun (University of Aberdeen)

Sarah Barry (University of Glasgow)

Frances Elliot (Scottish Government / NHS Fife)

David Garden (NHS Highland)

Fiona Ramsay (NHS Forth Valley)

Andrew Daly (NHS Greater Glasgow and Clyde)

Alisdair McDonald (NHS Lothian) – from January 2015 onwards

Pauline Craig (NHS Health Scotland) – from March 2014 onwards

Angela Campbell (Scottish Government)

Paudric Osborne (Scottish Government)

Ellen Lynch (Scottish Government) – up to August 2014

Tom Russon (Scottish Government) – from August 2014 to August 2015

Evan Williams (Scottish Government) – from January 2016 onwards

Roger Black (NHS ISD)

Petya Kindalova (NHS ISD) – from July 2014 to July 2015

Lynne Jarvis (NHS ISD) – from June 2015 onwards

Ahmed Mahmoud (NHS ISD) – up to June 2015

Peter Martin (NHS ISD) – from January 2016 onwards

Donna Mikolajczak (NHS ISD) – up to June 2015

Chris Mueller (NHS ISD) – from July 2015 to July 2016

Judith Stark (NHS ISD) – up to June 2015

Sarah Touati (NHS ISD) – from January 2015 onwards

Suzy Whoriskey (NHS ISD) – up to July 2014

## ANNEX B: URBAN-RURAL CLASSIFICATIONS USED IN THE REVIEW ANALYSIS

2013/14 Scottish Government Urban-Rural Classification – 8-fold version:

1. Large urban areas - settlements of 125,000 or more people.
2. Other urban areas - settlements of 10,000 to 124,999 people.
3. Accessible small towns - settlements of 3,000 and 9,999 people and within 30 minutes drive of a settlement of 10,000 or more.
4. Remote small towns - settlements of between 3,000 and 9,999 people and with a drive time of over 30 minutes to a settlement of 10,000 or more.
5. Very remote small towns - settlements of 3,000 and 9,999 people and with a drive time of over 60 minutes to a settlement of 10,000 or more.
6. Accessible rural - areas with a population of less than 3,000 people, and within a 30 minute drive time of a settlement of 10,000 or more.
7. Remote rural - areas with a population of less than 3,000 people, and with a drive time of over 30 minutes but less than 60 minutes to a settlement of 10,000 or more.
8. Very remote rural - areas with a population of less than 3,000 people, and with a drive time of over 60 minutes to a settlement of 10,000 or more.

2013/14 Scottish Government Urban-Rural Classification – 6-fold version:

1. Large urban areas - settlements of 125,000 or more people.
2. Other urban areas - settlements of 10,000 to 124,999 people.
3. Accessible small towns - settlements of 3,000 to 9,999 people and within 30 minutes drive of a settlement of 10,000 or more.
4. Remote small towns - settlements of 3,000 to 9,999 people and with a drive time of over 30 minutes to a settlement of 10,000 or more.
5. Accessible rural - areas with a population of less than 3,000 people, and within a 30 minute drive time of a settlement of 10,000 or more.
6. Remote rural - areas with a population of less than 3,000 people, and with a drive time of over 30 minutes to a settlement of 10,000 or more.

NHS Highland 4-fold classification, based on combining categories 1, 2 and 3 in the 6-fold classification above:

1. Urban areas - settlements of at least 10,000 people or at least 3,000 people within 30 min drive to a settlement of at least 10,000 people.
2. Accessible rural - areas with a population of less than 3,000 people, and within a 30 minute drive time of a settlement of 10,000 or more.
3. Remote small towns - settlements of 3,000 to 9,999 people and with a drive time of over 30 minutes to a settlement of 10,000 or more.
4. Remote rural - areas with a population of less than 3,000 people, and with a drive time of over 30 minutes to a settlement of 10,000 or more.

A 2-fold classification based on combining categories 2, 3 and 4 in the 4-fold classification above:

1. Urban areas - settlements of at least 10,000 people or at least 3,000 people within 30 min drive to a settlement of at least 10,000 people.
2. Rural areas - accessible rural areas, remote small towns and remote rural areas combined.

1. Whole time equivalent GP numbers are no longer available by GP practice. We therefore used GP headcount instead, although this may be an inferior indicator of GP supply. [↑](#footnote-ref-1)