**TAGRA ACUTE MLC SUBGROUP Tuesday 15th March 2016**

**INDICATOR SELECTION RESULTS – PART 2**

**1. Background and Summary**

Paper TAMLC40 presented results from the first stage of the Acute MLC indicator selection process – a list of ‘retained variables’ obtained by eliminating near-duplicates. The paper was discussed at the 11th meeting of the Subgroup. A number of actions were requested:

* Include ethnicity variables in the candidate variable list. Large numbers of zeros can be avoided either by combination of ethnic groups, or by computing the variable at intermediate zone level and applying the values to the contained data zones.
* Explore the possibility of distributing Outpatients activity across the other diagnostic groups based on specialty.
* Explore the possibility of including a variable to represent the presence of a prison within a data zone, so that the additional cost associated with prisoner Acute care can be modelled.

These actions have all been carried out and the results are reported in section 2 of this paper. Results from the second stage of the indicator selection process are presented in section 3: the variables retained from the first stage are reduced to a “restricted set” that together are able to explain almost as much variation in the cost ratios as the full retained set. In the third and final stage of the methodology, the most important few variables are identified – for individual diagnostic groups as well as for a “common” approach; these results are presented in section 4. In section 5 the options for the final index are analysed in terms of explanatory and predictive power. In section 6, the decisions required are laid out, and the relevant information is summarised and evaluated in terms of the TAGRA core criteria where possible.

Decisions from the Subgroup are required on:

* Exclusion of the Dementia variable, which has produced unexpected results
* How many indicators should be included in the final index / indices
* Common index vs specific indices for different diagnostic groups
* Retaining diagnostic groups vs ‘Whole Acute’ option

**2. First stage of indicator selection process: updates**

The methodology originally proposed in paper TAMLC36 for selecting the needs indicators is included in Annex A of the current paper, for reference (this has now been updated to include the final steps for diagnostic group-specific index options). In paper TAMLC40 the list of *retained* variables was presented after the first stage of the indicator selection process: elimination of near-duplicate variables.

Following discussion of paper TAMLC40, explorations were carried out around possible coding of Outpatient activity into diagnostic groups, inclusion of a simple indicator variable to account for prisons, and inclusion of various ethnicity variable options. The updates resulting from this extra work are presented in the following sub-sections.

**2.1 Outpatients diagnostic coding**

At the 11th meeting, it was suggested to use specialty information on Outpatient records (in SMR00) to assign a diagnostic group to each record, to avoid having Outpatients as a separate group.

Inpatients and Daycases (SMR01) data has been used to check if a clear mapping could be derived between specialty and diagnostic groups, which could then be applied to Outpatients (SMR00). This analysis showed that, for SMR01, the specialty given in the data is no clear guide to the diagnostic group.

An alternative would be to use a logical assignment of diagnostic category to SMR00 activity based on the specialty description. With this method, the majority of activity (82%) would be placed in the ‘Other’ category, which would not be likely to lead to improved predictions of cost as compared to analysing Outpatients separately.

Given the difficulty in either inferring a mapping between specialty and diagnostic group from the Inpatients/Daycases data or constructing such a mapping logically, there is no effective way of assigning diagnostic groups to outpatients. A decision was therefore taken to keep Outpatients as a separate group for the purposes of modelling healthcare need (unless the ‘Whole Acute’ option is eventually chosen by the Subgroup).

**2.2 Prison indicator variable**

Another suggestion from the 11th meeting was to explore the possibility of including a ‘dummy’ variable (i.e. indicator with value 0 or 1) to represent the presence of a prison within a data zone, rather than just excluding outlying prison data zones from the model, so that the additional cost associated with prisoner Acute care can be modelled and included in the target shares. A prison dummy variable was added to the reference model to test this idea; its coefficient was only significantly different from zero for Outpatients. This means 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 prison dummy variable has therefore been implemented only for Outpatients.

**2.3 Changes to retained variable set**

Some small changes to the retained variable list have now been made. Firstly, the DNA variable chosen previously was ‘DNA – ratio to population’. Expressing the DNA count as a ratio to the population potentially ran the risk of it acting as a proxy for total OP activity, since the more appointments there are per resident, the more DNAs there will be on average. This was not the intention behind the variable. We therefore exchanged this for the variant which expresses DNA count as a fraction of appointments, which should represent only the likelihood of a person missing a given appointment, not the relative frequency of appointments.

Secondly, ethnicity variables have been added. After the discussion at the 11th meeting, it was decided to include the following in the original list of candidate variables:

(1) A variable based on the count of all minority populations (including white minorities). This is not expected to show a strong relationship with healthcare use, but the results will be examined and further advice sought if there is found to be a strong relationship.

(2) A variable based on the count of all populations with worse than average health in the ScotStat report[[1]](#footnote-1). This comprises Gypsy/Traveller and Pakistani ethnic groups. The variable would need to be constructed at the more aggregated intermediate zone level to avoid large numbers of zeros, and the values applied to the contained data zones.

(3) Two variables based on the separate counts of Gypsy/Traveller and Pakistani populations, respectively. These variables also have to be constructed at intermediate zone level.

(4) A variable based on the count of all populations with better than average health in the ScotStat report. This is motivated by the observation that some of these populations (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 variables (2) or (3).

Near-duplicates analysis has been performed on these variables, and three ethnicity variables are retained to the next stage: both of the variables in (3) and the single variable in (4).

The updated list of retained variables is:

* All cause SMR <75
* Cancer SMR <75
* Heart SMR <75
* Other SMR <70
* LLTI – Yes (both)
* General health – bad or very bad
* Living alone ≥70

*Results from stage 1 of the process:* ***retained variable set.***

* 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
* Long-term sick and not seeking work
* Ethnic populations with better than average health
* Pakistani populations
* Gypsy/traveller populations

**3. Second stage of indicator selection process**

The second stage of the indicator selection process (see Annex A) aims to further reduce the variables retained from the first stage, to a “restricted set” of variables that together are able to explain almost as much variation in the cost ratios as the full retained set.

The procedures involved are as follows:

1. Use various regression techniques to iteratively select the most significant variables and reject the less powerful ones.
2. Perform principal component analysis, as an independent, alternative method to select the variables that best represent the retained set overall.
3. Check which variables chosen in step 2 were found to be significant in step 1 in the majority of diagnostic groups (4 or more). Those variables will form the common restricted variable set.

Section 3.1 describes the first step, section 3.2 relates to the second step, and section 3.3 presents the resulting restricted variable set.

**3.1 Iterative regressions**

In order to reduce the retained variable list down to a restricted set, we need some way of comparing the variables in terms of how well they explain the variation in cost. Ultimately we want the final indicators to explain as much variation as possible, so the restricted set must also be derived with that aim in mind. This is complicated by the fact that there are different degrees of inter-correlation between the variables, which means that the best set of *individual* variables may well not be the best possible set in *combination*.

This can be achieved using iterative regression methods (stepwise regression, forward regression and backward regression); in these methods, variables are added and/or removed at each iteration in such a way as to eventually arrive at the best-performing subset of variables.

There is always a trade-off between the increase in R2 that can be obtained by adding additional variables, and the desire to shrink the variable list. Reducing the variable list is necessary not just for practical reasons but because there is also a risk of ‘over-fitting’, whereby variables that increase the R2 based on the current data set do not actually improve the *predictive* power of the model. The iterative regression methods introduce a ‘penalty’ for each additional variable that is added[[2]](#footnote-2), to mitigate the over-fitting problem.

Regularisation methods (lasso and elastic net) for regression work in a similar way. In these methods the penalty is set up to ‘shrink’ the estimates of the regression coefficients towards zero, resulting in elimination of variables.

Each of these types of regressions was carried out with the cost ratios and the 15 retained variables (with the supply model always included, and for Outpatients, the prison dummy always included). Using a variety of methods allows greater confidence in the robustness of the results. Table 1 shows which variables were selected, by diagnostic group, using the different methods. (The results from elastic net regression were identical to those from lasso regression, so those results are omitted.)

*Table 1: Variables that appear as significant using stepwise / forward / backward / lasso regression, respectively. Empty cells indicate that none of the methods found the variable to be significant.*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Whole Acute** | **Cancer** | **Heart** | **Digestive** | **Injury** | **Other** | **Respiratory** | **Outpatients** |
| All cause SMR <75 | ✓✓✓✓ | ✓✓✓🗶 | 🗶🗶🗶✓ | ✓✓✓✓ | 🗶🗶🗶✓ | 🗶✓🗶✓ | ✓✓✓✓ | ✓✓✓🗶 |
| Cancer SMR <75 | ✓✓✓✓ | ✓✓✓✓ |  | ✓✓✓🗶 |  |  |  | ✓✓✓✓ |
| Heart SMR <75 |  |  | ✓✓✓✓ | ✓✓✓🗶 |  |  |  |  |
| Other SMR <70 |  |  |  | ✓✓✓🗶 |  | ✓🗶✓🗶 |  |  |
| LLTI | ✓✓✓✓ |  | ✓✓✓✓ | 🗶✓🗶✓ | ✓✓✓✓ | ✓✓✓✓ | 🗶✓🗶✓ | 🗶🗶🗶✓ |
| General health | ✓✓✓✓ |  | 🗶🗶🗶✓ | ✓✓✓✓ | 🗶🗶🗶✓ | 🗶🗶🗶✓ | ✓✓✓✓ | ✓✓✓✓ |
| Living alone ≥70 | ✓✓✓✓ | ✓✓✓🗶 | ✓✓✓🗶 |  |  | ✓✓✓🗶 | ✓✓✓🗶 |  |
| Living alone ≥90 | ✓✓✓✓ |  | ✓✓✓🗶 |  |  |  |  |  |
| Unpaid care ≥ 20 hours | ✓✓✓✓ |  |  | ✓✓✓✓ |  | ✓✓✓✓ | ✓✓✓✓ | ✓✓✓✓ |
| Education – level 2 and below |  |  |  |  | ✓🗶✓🗶 |  |  | ✓✓✓✓ |
| DNA counts | ✓✓✓✓ |  | ✓✓✓✓ | ✓✓✓✓ | ✓✓✓✓ | ✓✓✓✓ | ✓✓✓✓ | ✓✓✓✓ |
| Low birth weight births |  |  |  |  |  | ✓✓✓✓ |  |  |
| Dementia | ✓✓✓✓ | ✓✓✓✓ | ✓✓✓✓ | ✓✓✓✓ | ✓✓✓✓ | ✓✓✓✓ | ✓✓✓✓ | ✓✓✓✓ |
| High resource individuals | ✓✓✓✓ | ✓✓✓✓ | ✓✓✓✓ | ✓✓✓✓ | ✓✓✓✓ | ✓✓✓✓ | ✓✓✓✓ | ✓✓✓✓ |
| Long-term sick and not seeking work |  |  |  |  | ✓🗶✓🗶 |  |  | ✓✓✓🗶 |
| Ethnic populations with better health  | ✓✓✓✓ |  |  | ✓✓✓✓ |  | ✓✓✓✓ | ✓✓✓🗶 | ✓✓✓✓ |
| Pakistani population |  |  |  | ✓✓✓🗶 |  |  |  | ✓✓✓🗶 |
| Gypsy/Traveller population |  |  |  |  |  |  |  | ✓✓✓✓ |

Comparing the different iterative regression techniques, the variables deemed significant (Table 1) are mostly the same across all methods, and the differences in the final adjusted R2 are minor. Therefore, to be consistent with the outline in Annex A, the outputs from the *stepwise* regression are used.

**3.2 Principal components** **analysis**

Principal component analysis (or factor analysis) reduces a set of variables into a smaller number of ‘principal components’ that will account for most of the variance in the original variables. Each principal component is essentially a composite variable consisting of each of the input variables multiplied by a weighting. The weightings reflect the association between each of the original variables and each of the components; higher weightings indicate that the variable plays an important role in explaining the overall variance.

This technique is useful as an independent, alternative approach to selecting the variables that best represent the retained set overall. The principal components themselves are not useful to us as variables since they are not transparent, or do not have a clear meaning. However, they can be used as a guide to which of the original variables are the most important.

Table 2 shows the five principal components that have been produced by the method, along with the weighting of each of the retained variables in each component. Weightings exceeding an absolute value of 0.40 are highlighted in bold.

For each component, the four variables with the highest weightings are provisionally selected. This produces a provisional restricted set. The variables in this list that appear significant across most diagnostic groups in the iterative regressions (Table 1) will form the restricted variable set.

*Table 2: Principal component analysis results*

|  |  |
| --- | --- |
|   | **Component** |
| **1** | **2** | **3** | **4** | **5** |
| All cause SMR<75 | **0.919** | 0.148 | 0.054 | 0.043 | -0.012 |
| Cancer SMR<75 | **0.651** | 0.049 | 0.036 | 0.050 | 0.021 |
| Heart SMR<75 | **0.717** | 0.092 | 0.017 | 0.028 | -0.005 |
| Other SMR<70 | **0.578** | 0.233 | 0.077 | 0.175 | -0.047 |
| LLTI | **0.879** | 0.056 | 0.018 | -0.318 | 0.043 |
| General health | **0.881** | -0.020 | -0.004 | -0.285 | -0.023 |
| Living alone>=70 | 0.017 | **-0.646** | 0.187 | 0.263 | -0.263 |
| Living alone>=90 | 0.036 | -0.267 | 0.002 | -0.137 | **0.521** |
| Unpaid care>=20 | **0.573** | -0.018 | 0.060 | **-0.624** | 0.097 |
| Education | -0.035 | 0.141 | **0.854** | -0.280 | 0.132 |
| DNA counts | **0.818** | -0.269 | -0.058 | -0.052 | -0.160 |
| Low birth weight births | 0.190 | -0.002 | -0.029 | -0.076 | 0.037 |
| Dementia | -0.029 | **0.794** | 0.084 | 0.067 | -0.176 |
| High resource individuals | 0.369 | **0.765** | 0.111 | -0.129 | 0.007 |
| Long-term sick and not seeking work | 0.062 | -0.073 | **0.917** | 0.075 | -0.039 |
| Pakistani population | 0.160 | -0.221 | -0.074 | 0.113 | **-0.630** |
| Gypsy/Traveller population | 0.172 | -0.003 | 0.000 | 0.448 | **0.588** |
| Ethnic populations with better health | -0.106 | -0.169 | -0.128 | **0.788** | -0.081 |

**3.3 Restricted variable set**

Considering the results of both the iterative regressions in section 3.1 and the principal components analysis in section 3.2, the common restricted set consists of:

* LLTI

*Results from stage 2 of the process:* ***common restricted variable set.***

* High resource individuals
* DNA counts
* Ethnic populations with better health
* All cause SMR<75
* Dementia
* Unpaid care ≥ 20 hours
* Living alone ≥70

Table 3 shows the R2 resulting from the common restricted set, in comparison with the R2 obtained from using the full list of retained variables. It is clear that R2 has not decreased very much by eliminating some variables, except for the diagnostic group ‘Cancer’. The R2 obtained using the current Acute MLC indicators, and using the supply model only, are also shown for reference.

*Table 3: Adjusted R2 for the common restricted variable set, compared with other models*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Whole Acute** | **Cancer** | **Heart** | **Digestive** | **Injury** | **Other** | **Respiratory** | **Outpatients** |
| Retained variable set + supply model | 71.4% | 19.1% | 25.5% | 44.1% | 31.4% | 54.5% | 43.7% | 56.1% |
| Common restricted variable set+ supply model | 71.1% | 15.0% | 25.2% | 43.5% | 31.0% | 54.3% | 42.9% | 55.4% |
| Current indicators (LLTI + All-cause SMR) + supply model | 61.8% | 11.2% | 22.0% | 39.8% | 27.3% | 47.2% | 39.2% | 51.1% |
| Supply model only | 18.8% | 6.9% | 3.0% | 17.6% | 3.4% | 15.3% | 10.5% | 40.5% |

**4. Third stage of indicator selection process**

The third and final stage of the indicator selection process (see Annex A) is to identify the few best-performing indicators that will form the proposed final index. The ‘top’ 1, 2, 3 and 4 variables are identified, and at this stage two approaches are used – the ‘common’ approach (as for stages 1 and 2), and the ‘specific’ approach in which each diagnostic group has its own top 4. Section 4.1 describes the methodology for doing this, and sections 4.2 and 4.3 present the results.

**4.1 Methodology for selecting the index components**

For the ‘common’ index approach, having reduced the number of variables down to 8 using iterative regressions and principal components analysis, we now need to find the best-performing subsets of those.

Table 4 shows the R2 obtained from using each retained variable as the sole indicator in a regression. As noted in section 3.1, these values alone cannot determine the best combinations of variables, as this depends on the inter-correlations between the variables, which could vary considerably between different pairs of variables. For this reason, the methodology (Annex A) specifies the use of *coefficients* to rank the variables instead. A multiple regression is carried out using the common restricted set of 8 variables, along with the supply model, and the coefficients of the different variables are compared. The variables are ranked in order of highest to lowest coefficients. (The order may differ between diagnostic groups but the overall ranking is based on the majority of diagnostic groups.)

For constructing the *specific* top 4 for each diagnostic group, we must go back to the retained variable list (page 3) since it is possible that the common restricted set of 8 variables has excluded some variable that would work well for a specific diagnostic group. The top 1 is simply taken to be the variable with the highest individual R2 in Table 4, and then for the top 2, all possible pairs of variables are tested and the pair with the highest R2 is selected. The same exhaustive testing is carried out for the top 3 and the top 4.

Finally, the specific top 4 indicators for the different diagnostic groups are examined and the common top 4 are checked against these – variables that commonly appear in the specific top 4s would be expected to appear in the common top 4 also, so this provides further checking and validation of the ‘coefficients criterion’ for selecting the common top indicators.

*Table 4: R2 for each of the variables in the retained variable list, when it is used as the only indicator (along with the supply model and prison dummy). Highest R2 is shown in bold.*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Whole Acute** | **Cancer** | **Heart** | **Digestive** | **Injury** | **Other** | **Respiratory** | **Outpatients** |
| All cause SMR<75 | 48.1% | 11.1% | 15.8% | 32.5% | 19.5% | 36.0% | 31.1% | 45.7% |
| Cancer SMR <75 | 33.1% | **15.3%** | 8.2% | 23.9% | 9.3% | 24.0% | 18.5% | 43.7% |
| Heart SMR <75 | 34.3% | 8.0% | 11.3% | 24.8% | 12.5% | 26.9% | 21.4% | 43.1% |
| Other SMR <70 | 27.1% | 7.3% | 7.0% | 21.2% | 7.7% | 22.4% | 16.8% | 41.7% |
| LLTI | **60.8%** | 9.7% | **21.6%** | **39.3%** | **26.8%** | **46.7%** | 38.2% | **51.1%** |
| General health | 57.8% | 9.4% | 20.6% | 38.3% | 24.9% | 43.5% | **38.5%** | 50.4% |
| Living alone ≥70 | 19.5% | 6.9% | 3.1% | 18.1% | 4.1% | 16.0% | 10.7% | 40.9% |
| Living alone ≥90 | 19.0% | 7.1% | 3.3% | 17.7% | 3.5% | 15.4% | 10.6% | 40.6% |
| Unpaid care ≥20 | 43.2% | 8.5% | 12.3% | 30.4% | 13.0% | 33.8% | 25.8% | 51.0% |
| Education | 20.5% | 7.0% | 3.4% | 18.2% | 4.2% | 16.7% | 11.1% | 41.8% |
| DNA counts | 44.7% | 8.4% | 16.2% | 31.2% | 19.1% | 33.9% | 29.5% | 45.6% |
| Low birth weight births | 20.0% | 7.0% | 3.4% | 18.2% | 3.8% | 16.8% | 11.3% | 40.6% |
| Dementia  | 18.8% | 7.4% | 3.2% | 17.6% | 3.5% | 15.3% | 10.5% | 40.7% |
| High resource individuals | 38.5% | 9.5% | 9.6% | 26.5% | 15.0% | 31.2% | 21.1% | 44.3% |
| Long term sick and not seeking work | 19.2% | 7.0% | 3.2% | 17.7% | 3.5% | 15.5% | 10.9% | 40.6% |
| Ethnic populations with better health | 25.4% | 7.3% | 4.5% | 21.5% | 5.1% | 21.0% | 13.9% | 44.6% |
| Pakistani populations | 19.9% | 7.1% | 3.4% | 18.6% | 4.1% | 16.0% | 11.0% | 40.9% |
| Gypsy/Traveller populations | 19.4% | 7.0% | 3.4% | 17.9% | 4.3% | 15.8% | 11.2% | 40.5% |

**4.2 Exclusion of Dementia**

Table 5 shows the standardised coefficients of the common restricted set, obtained from a multiple linear regression including all 8 variables. For ‘Dementia’ and ‘Ethnic group populations with better health’, the coefficients are negative across all diagnostic groups. In the latter case, this is expected, since we are using the proportion of the population belonging to any ethnic group with reported better than average health – we would expect this to be negatively correlated with cost. However, for Dementia, relatively high negative coefficients are unexpected and difficult to interpret.

While high coefficients in Table 5 mostly correspond to high R2 values in Table 4, Dementia is an exception; the R2 values for Dementia when used as the sole indicator are not particularly high.

Figure 1 shows a scatter plot of Whole Acute cost ratios against the Dementia variable. This shows that there is not a clear linear relationship with a negative slope; the best-fit line has a negative slope because there is a large – and skewed – spread of cost ratio values, particularly at low values of the variable.

It is possible that Dementia tends to afflict otherwise healthy individuals who have lived to an old age, and so is not a good predictor of need for Acute services.

*Table 5: Standardised coefficients for the common restricted variable set, obtained from a multiple regression. Values greater than ±0.100 are in bold.*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Whole Acute** | **Cancer** | **Heart** | **Digestive** | **Injury** | **Other** | **Respiratory** | **Outpatients** |
| All cause SMR<75 | 0.093 | **0.170** | 0.050 | 0.068 | 0.035 | 0.048 | **0.108** | -0.013 |
| LLTI | **0.229** | -0.060 | **0.197** | **0.180** | **0.251** | **0.218** | **0.184** | **0.116** |
| Unpaid care ≥20 | **0.105** | 0.033 | 0.034 | 0.069 | -0.018 | 0.088 | 0.082 | **0.167** |
| Living alone ≥70 | 0.046 | 0.041 | 0.038 | 0.016 | 0.005 | 0.040 | 0.051 | 0.004 |
| DNA counts | **0.207** | -0.021 | **0.154** | **0.157** | **0.194** | **0.180** | **0.197** | 0.059 |
| Dementia  | **-0.244** | **-0.226** | **-0.170** | **-0.144** | **-0.119** | **-0.187** | **-0.118** | **-0.139** |
| High resource individuals | **0.410** | **0.259** | **0.245** | **0.235** | **0.280** | **0.357** | **0.246** | **0.166** |
| Ethnic populations with better health | **-0.110** | -0.007 | -0.030 | **-0.104** | -0.027 | **-0.113** | -0.079 | **-0.120** |



*Figure 1: Scatter plot of Whole Acute cost ratios against Dementia.*

**Q: AST propose to exclude Dementia from the top 4, since its performance (in terms of R2) is relatively poor, the scatter plot suggests that there is no real relationship with cost, and the slope of the fitted line is unexpectedly negative. The Subgroup is asked to approve this suggestion.** Table 6 evaluates the decision to exclude Dementia against the relevant TAGRA core criteria.

*Table 6. TAGRA core criteria evaluation for excluding Dementia.*

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Decision:** | **Equity** | **Practicality** | **Transparency** | **Objectivity** | **Avoiding Perverse Incentives** | **Relevance** | **Stability** | **Responsive-ness** | **Face Validity** |
| **Exclude Dementia from Top 4** |  |  |  | + Dementia R2 performance is relatively poor+ Scatter plots reveal there is no real relationship between the Dementia variable and cost |  |  |  | + The Dementia variable is not a good predictor of cost (low R2) | + Negative slope is counter-intuitive |

**4.3 The top 4 indicators**

After excluding the Dementia variable, the common top 4 – derived using the coefficients criterion – were checked against the specific top indicators. A discrepancy was found for the top 2. The common top 2 consisted of LLTI and DNA; however, LLTI and HRI are the specific top 2 for Whole Acute, Heart, Digestive, Injury and Other. It therefore seemed more reasonable to use LLTI and HRI as the common top 2 as well. A check of the R2 confirmed that LLTI and HRI perform better over most diagnostic groups than LLTI and DNA. (Additionally, in regressions on the “common top 2” indicators, it was found that DNA had a negative coefficient for both Cancer and Outpatients; when it is replaced by HRI this does not happen.)

The resulting top 4 indicators are shown in Table 7. For Whole Acute, Heart, Injury and Other, the specific top 3 is the same set of variables as the common top 3. Digestive and Respiratory have two out of their specific top 3 in common with the common set. General Health sometimes appears in place of LLTI (this makes sense as they are conceptually similar). Cancer, on the other hand, has a variable appearing as its top 1 that was not even in the common restricted set: Cancer SMR <75.

*Table 7: Top 4 indicators*

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Common top 4** | **Whole Acute top 4** | **Cancer specific top 4** |
| **Top 1****Top 2****Top 3****Top 4** | LLTI LLTI, HRI LLTI, HRI, DNALLTI, HRI, DNA, Ethnic populations with better health | LLTI LLTI, HRILLTI, HRI, DNALLTI, HRI, DNA,General health | Cancer SMRCancer SMR, HRICancer SMR, HRI, All-cause SMRCancer SMR, HRI, All-cause SMR, Living alone |
|  | **Heart specific top 4** | **Digestive specific top 4** | **Injury specific top 4** |
| **Top 1****Top 2****Top 3****Top 4** | LLTI LLTI, HRILLTI, HRI, DNALLTI, HRI, DNA, General health | LLTI LLTI, HRIGeneral health, HRI, DNA General health, HRI, DNA, Ethnic populations with better health | LLTI LLTI, HRILLTI, HRI, DNALLTI, DNA, HRI, General health |
|  | **Other specific top 4** | **Respiratory specific top 4** | **Outpatients specific top 4** |
| **Top 1****Top 2****Top 3****Top 4** | LLTI LLTI, HRILLTI, HRI, DNALLTI, HRI, DNA, Ethnic populations with better health | General health General health, HRIGeneral health, HRI, DNAGeneral health, HRI, DNA, Ethnic populations with better health | LLTILLTI, Unpaid careUnpaid care, General health, Ethnic populations with better healthUnpaid care, General health, Ethnic populations with better health, HRI |

*Results from stage 3 of the process:* ***common and specific ‘top 4’ indicators.***

**5. Analysis of the index options**

Having derived the top 1, 2, 3 and 4 indicators in section 4, in this section the performance of the different index options is examined. Section 5.1 looks at how well the different options explain variation in the 3-year cost ratios, and section 5.2 looks at how well the different options predict the future costs.

**5.1 Explanatory power: adjusted R2**

Table 8 shows the adjusted R2 values from the regressions. The adjusted R2 values using the current Acute indicators of need – LLTI and All-cause SMR – are also shown, for comparison. Higher R2 values are produced using the specific top 2 variables, and the common top 2 (with the exception of Cancer), compared with the current indicators, but the difference is quite small.

As noted in section 3.1, there is always a trade-off between the increase in R2 that can be obtained by including more variables, and the desire to shrink the variable list. However, there is hardly any reduction in R2 in going from a 4-indicator model to a 3-indicator model, for most diagnostic groups. Further reducing the number of indicators to 2 and again to 1 results in a slightly more noticeable reduction in R2, but the reduction is still fairly modest.

For the diagnostic group Cancer, R2 is consistently higher when using its specific top 1, 2, 3 or 4 variables than when using the common set. For other diagnostic groups, there are no substantial differences in R2 between the common and specific indicator sets.

*Table 8: Adjusted R2 for the top 1, 2, 3 and 4 variables from the restricted sets*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Whole Acute** | **Cancer** | **Heart** | **Digestive** | **Injury** | **Other** | **Respiratory** | **Outpatients** |
| R2 using common top 4 variables  | 66.6% | 10.7% | 23.3% | 42.0% | 30.2% | 51.7% | 41.2% | 53.0% |
| R2 using specific top 4 variables | 65.7% | 16.4% | 23.3% | 42.1% | 30.3% | 51.7% | 42.2% | 54.3% |
| R2 using common top 3 variables | 65.4% | 10.7% | 23.2% | 41.0% | 30.2% | 50.6% | 40.7% | 51.5% |
| R2 using specific top 3 variables | 65.4% | 16.2% | 23.2% | 40.9% | 30.2% | 50.6% | 41.5% | 53.9% |
| R2 using common top 2 variables | 64.2% | 10.6% | 22.2% | 40.5% | 28.9% | 49.8% | 39.4% | 51.5% |
| R2 using specific top 2 variables | 64.2% | 16.1% | 22.2% | 40.5% | 28.9% | 49.8% | 40.3% | 53.0% |
| R2 using common top 1 variables | 60.8% | 9.7% | 21.6% | 39.3% | 26.8% | 46.7% | 38.2% | 51.1% |
| R2 using specific top 1 variables | 60.8% | 15.3% | 21.6% | 39.3% | 26.8% | 46.7% | 38.5% | 51.1% |
| ***Current indicators (LLTI + All-cause SMR)*** | ***61.8%*** | ***11.2%*** | ***22.0%*** | ***39.8%*** | ***27.3%*** | ***47.2%*** | ***39.2%*** | ***51.1%*** |

In the NRAC formula, the indicators of need are transformed into a ‘needs index’ (as the sum of the Z-scores of the indicators). Regression is then done on the needs index along with the supply model. Table 9 shows the adjusted R2 values from ‘index’ regressions based on the top 1, 2, 3 and 4 indicators. The main observations are:

* The values are generally slightly lower than those in Table 8.
* There is no benefit to using 4 indicators over 3 and in most cases the R2 is actually *lower* for a 4-indicator index.
* Again, only for Cancer is the R2 consistently higher when using the specific top 1, 2, 3 or 4 variables.
* The R2 values for the reference model are now sometimes higher than those for the common and specific top 2 indicators.

*Table 9: Adjusted R2 for the top 1, 2, 3 and 4 variables combined into an index, in comparison with the current reference model*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Whole Acute** | **Cancer** | **Heart** | **Digestive** | **Injury** | **Other** | **Respiratory** | **Outpatients** |
| R2 using common top 4 index  | 52.9% | 9.7% | 19.0% | 34.1% | 25.5% | 40.5% | 33.3% | 46.4% |
| R2 using specific top 4 index | 65.5% | 14.2% | 23.3% | 33.6% | 30.1% | 40.5% | 33.2% | 49.2% |
| R2 using common top 3 index | 64.9% | 10.6% | 22.7% | 40.6% | 30.1% | 50.2% | 40.1% | 50.5% |
| R2 using specific top 3 index | 64.9% | 14.5% | 22.7% | 40.6% | 30.1% | 50.2% | 40.8% | 47.2% |
| R2 using common top 2 index | 60.2% | 10.6% | 19.4% | 37.8% | 27.0% | 47.3% | 35.6% | 49.9% |
| R2 using specific top 2 index | 60.2% | 15.1% | 19.4% | 37.8% | 27.0% | 47.3% | 36.4% | 53.0% |
| R2 using common top 1 index | 60.8% | 9.7% | 21.6% | 39.3% | 26.8% | 46.7% | 38.2% | 51.1% |
| R2 using specific top 1 index | 60.8% | 15.3% | 21.6% | 39.3% | 26.8% | 46.7% | 38.5% | 51.1% |
| ***Reference model*** | ***59.6%*** | ***10.9%*** | ***21.0%*** | ***38.6%*** | ***26.0%*** | ***45.1%*** | ***38.2%*** | ***49.3%*** |

**5.2 Predictive power: RSS**

*Predictive* power is arguably more important than *explanatory* power, since the MLC adjustment is used to predict cost ratios in the year of allocation. To evaluate the models in predictive mode, predicted cost ratios are generated. These predictions are then compared with a 1-year cost ratio based on 2014/15 data. The 2014/15 cost ratio represents the ‘future’ observation which the model would be trying to predict.

Predicted cost ratios are calculated in the same way as in the NRAC formula: using only the indicators of need, transformed into a ‘needs index’ (as the sum of the Z-scores of the indicators). The coefficient of the needs index is obtained through a regression including the supply model, but the supply variables are not used in the prediction. In the case of Outpatients, the prison dummy is included in both the regression and the prediction.

A useful measure for the comparison of predictions with observations is the residual sum of squares (RSS): this is the sum of the squared differences between the predictions and the observations. Low RSS values indicate that the observations are relatively close to the predictions.

The RSS values are given in Table 10. Generally, the RSS either stays much the same or even *decreases* in going from a 4-variable to a 3-variable model. This fits with the increase in R2 observed in Table 9, and implies that using 3 variables results in as good or better predictions than using 4 variables, at least for the 2014/15 costs. Further reducing the model to 2 or 1 variable results in some small changes to the RSS, both positive and negative.

*Table 10: RSS obtained from comparing predictions derived from the top 1, 2, 3 and 4 variables from the restricted sets with the 2014/15 cost ratios. Lower values indicate the predictions are closer to the observed values.*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Whole Acute** | **Cancer** | **Heart** | **Digestive** | **Injury** | **Other** | **Respiratory** | **Outpatients** |
| RSS using common top 4 index | 408 | 3476 | 4063 | 2360 | 2915 | 914 | 3663 | 424 |
| RSS using specific top 4 index | 344 | 3516 | 4029 | 2363 | 2868 | 914 | 3666 | 427 |
| RSS using common top 3 index | 354 | 3472 | 4045 | 2197 | 2875 | 811 | 3454 | 412 |
| RSS using specific top 3 index | 354 | 3542 | 4045 | 2194 | 2875 | 811 | 3443 | 427 |
| RSS using common top 2 index | 408 | 3511 | 4112 | 2274 | 2936 | 865 | 3646 | 428 |
| RSS using specific top 2 index | 408 | 3580 | 4112 | 2274 | 2936 | 865 | 3614 | 416 |
| RSS using common top 1 index | 355 | 3467 | 4015 | 2160 | 2861 | 821 | 3499 | 417 |
| RSS using specific top 1 index | 355 | 3554 | 4015 | 2160 | 2861 | 821 | 3501 | 417 |
| ***RSS using reference model*** | ***368*** | ***3471*** | ***4028*** | ***2203*** | ***2868*** | ***840*** | ***3547*** | ***422*** |

It is striking that despite the higher R2, the specific variables for Cancer do not perform better in predictive mode than the common set; they actually perform slightly *worse* (although the difference is quite small).

Out of all the variables, Cancer SMR <75 is most clearly related to the dependent variable: 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. (The Cancer SMR variable was based on deaths between April 2009 and March 2014; the cost ratios use costs and activity from April 2011 to March 2014.) 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. The higher R2 for the specific model for Cancer is potentially misleading.

It is also striking that the RSS for the reference model is actually lower (i.e. predictions are better) compared to the common or specific ‘top 2’ models. This could possibly be due to a similar effect with HRI as with Cancer SMR: it is also fairly closely related to the dependent variable, as both are based on actual cost. However, many or most of the high resource individuals counted in the variable can be expected to go on living into the year of allocation and are likely to continue to use a large amount of resource, so there is likely to be less of an issue with HRI than with Cancer SMR.

**6. Decisions required**

In this section, the decisions required are laid out, and the relevant information from the preceding sections is summarised and evaluated in terms of the TAGRA core criteria where possible.

The decisions required are:

1. How many indicators should be included in the final model? Are there non-statistical reasons to prefer some variables over others? (Section 6.1)
2. Should we use a specific index for each diagnostic group, or a common index for all groups? (Section 6.2)
3. Should we retain the diagnostic groups, or apply the MLC adjustment to the Acute care programme as a whole (‘Whole Acute’ option)? (Section 6.3)

**6.1 Which components should be included in the index?**

Table 11 shows, for each of the variables that appear in the ‘top 4’ lists, full details of how the variable was calculated and the original rationale for its inclusion in the candidate variable list.

In section 5.1, no increase in explanatory power was found when using 4 indicators compared to 3; in most cases the R2 is actually *lower* for 4 indicators when looking at a combined index. In section 5.2 it was also found that 4 indicators offered no increase in predictive power over 3 indicators, when predictions were compared with the 2014/15 cost ratios. However, while it seems 4 components might be too many to include in an index, the analysis did not clearly favour the use of 1, 2 or 3 indicators.

As discussed in section 5.2, Cancer SMR <75 and HRI are more closely related to the dependent variable – the cost ratios – than the other variables. For Cancer SMR <75, this seems to result in significantly higher levels of *explanation* (for the past cost ratios used to fit the model) than *prediction* (for future costs). Regarding HRI, it did seem that the current reference model (LLTI + All-cause SMR) produced slightly better predictions than the top 2 (LLTI + HRI), which is possibly a sign of too close a dependence between the HRI variable and the cost ratios used to fit the model.

*Table 11: Variables appearing in the ‘top 4’ sets, with full details and rationale*

|  |  |
| --- | --- |
| **Variable** | **Details** |
| Limiting long-term illness  | Number of respondents answering ‘Yes’ (including both ‘a little’ and ‘a lot’) in Census 2011 question 21 – standardised by age and sex using 2011 MYE population. Included as a measure of the prevalence of serious morbidity. |
| High Resource Individual (HRI) counts | Count of individuals belonging to the group of ~100,000 highest resource-users that account for 50% of the total resource[[3]](#footnote-3). 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). Included as a measure of the prevalence of long-term conditions resulting in high healthcare costs.  |
| Did Not Attend (DNA) counts | Based on 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. Included because of its theoretical link to deprivation. |
| Ethnic populations with better health | Data from Census 2011; count of all populations with better than average self-reported health (i.e. all ethnic groups except White Scottish, Bangladeshi, Pakistani, and Gypsy/Traveller) – expressed as a simple fraction of 2011 MYE population. Included in order to differentiate between the resource needs of different ethnic groups. |
| All cause SMR <75 | Age-sex standardised mortality ratios with all causes of death. Calculated using 5 financial years’ GRO death records (09/10 – 13/14) and average population over middle 3 years (MYEs – 2011, 2012, 2013). Included as a measure of the prevalence of ‘premature’ death. |
| Cancer SMR <75 | As above, but cause of death selected using ICD10 codes C00--D48. |
| Dementia prescriptions | Number of patients receiving prescriptions in BNF section 4.11, based on 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). Included as an attempt to measure the prevalence of dementia, and to test whether it is associated with higher use of Acute services. |
| Older people living alone – 70 and over | Data from Census 2011 – standardised by age and sex using 2011 MYE population. Included in anticipation that older people living alone may access Acute services more frequently or have longer stays in hospital. |
| Unpaid care – 20 hours or more | Number of respondents indicating that they give 20 or more hours of unpaid care, in Census 2011 question 9 – standardised by age and sex using 2011 MYE population. Included to test whether caregiving is associated with high need for Acute services. |
| General health | Number of respondents answering ‘Bad’ or ‘Very bad’ in Census 2011 question 19 – standardised by age and sex using 2011 MYE population. Included as a measure of morbidity. |

Additionally, there may be non-statistical reasons to prefer some variables: the 4th ‘common’ variable is ethnicity, which may be preferred for inclusion on grounds of equity.

A perverse incentive could be perceived to be created by the use of a variable such as HRI that relates so directly to healthcare expenditure. On the other hand, around 60% of the individuals identified as HRIs are aged 65 or over, so this variable does support the requirement to capture the needs of the older population.

DNA rates were included because of their theoretical link to deprivation. However, arguably, there could be a perverse incentive if this variable were used in the Acute MLC adjustment: Boards may be less motivated to reduce their DNA rates if this would lower their predicted resource need.

**Q: The Subgroup is asked to consider the question of how many (and which) indicators to include in the final index(es).** Table 12 examines the arguments in terms of the TAGRA core criteria.

*Table 12. TAGRA core criteria evaluation for different variable combinations.*

|  |  |
| --- | --- |
| **Equity** | Equity would perhaps favour inclusion of the 4th ‘common’ variable, ethnicity  |
| **Practicality** | A model with fewer indicators would be more practical to implement and maintain |
| **Transparency** |  |
| **Objectivity** | 3 indicators perform just as well or better in an index than 4Cancer SMR <75 is poorer at predicting future cancer costs than at explaining past cancer costsLLTI + HRI performs better in regressions than LLTI + All-cause SMR, but worse in predictive mode -> possibly HRI is also poorer at predicting future costs than at explaining past costs |
| **Avoiding Perverse Incentives** | HRI and DNA could be seen as setting up perverse incentives (to concentrate resources on individuals; to maintain a high DNA rate) |
| **Relevance** | All variables have a strong theoretical/intuitive link to healthcare need Around 60% of the individuals identified as HRIs are aged 65 or over, so this variable could help to fulfil the requirement to capture the needs of the older populationDNA is perhaps a stronger proxy measure of deprivation than any of the other variables |
| **Stability** |  |
| **Responsiveness** | Around 60% of the individuals identified as HRIs are aged 65 or over, so this variable could help to fulfil the requirement to capture the needs of the older populationDNA is a proxy for deprivation |
| **Face Validity** | DNA would need careful presentation to not be mis-interpreted |

**6.2 Specific vs common indicator sets**

The specific top 4 indicators tend to have some overlap with the common set (Table 7). For Whole Acute, Heart, Injury and Other, the specific top 3 is the same as the common top 3, and Digestive and Respiratory have two out of their specific top 3 in common with the common set. The explanatory or predictive power is not consistently higher using the specific indicator set, for any of these cases. With such similar results between the ‘specific’ and ‘common’ approaches, we cannot be confident that the specific indicators will perform better over the long term and justify the extra complexity.

Cancer has a variable appearing as its top 1 that was not even in the common restricted set: Cancer SMR <75. Tables 8 and 9 showed that for Cancer, R2 is consistently higher when using its specific top 1, 2, 3 or 4 variables than when using the common set. However, in Table 10, the RSS values indicate that the specific variables for Cancer do not perform better in predictive mode than the common set; they actually perform slightly *worse* (although the difference is quite small). This could be because, as noted in section 5.2, the Cancer SMR <75 variable is closely related to the past Cancer cost ratios (due to temporally overlapping data) and may be less closely related to *future* cost ratios. It is also worth noting that, while carrying out the regressions for Table 8, All-cause SMR <75 was found to have a negative coefficient in the ‘specific’ regressions for Cancer; this is counter-intuitive, and is a further argument against using the specific indicators for Cancer.

The only other group for which the specific variables may be perceived to offer an improvement over the common-index approach is Outpatients. The specific variables for Outpatients do not have much overlap with the common set (Table 7); Unpaid Care (20 hours or more) is a strong variable for Outpatients and appears in the top 2, 3 and 4. The R2 using the specific set is mostly a little higher than for the common set (Tables 8 and 9). However, the RSS obtained by comparing predictions with 2014/15 cost ratios (Table 10) shows a more mixed result and does not clearly favour one variable set over the other.

**Q: AST propose that the ‘common index’ approach is retained. The Subgroup is asked to decide whether there is justification to use specific indices for some, or all, diagnostic groups.** Table 13 evaluates the arguments using the TAGRA core criteria.

*Table 13. TAGRA core criteria evaluation for common vs specific index approaches.*

|  |  |
| --- | --- |
|  | **Arguments for and against (+/-) diagnostic group-specific indices:** |
| **Equity** |  |
| **Practicality** | - It is more practical to use a common index |
| **Transparency** |  |
| **Objectivity** | - Specific variables do not generally seem to result in significantly better predictions |
| **Avoiding Perverse Incentives** |  |
| **Relevance** | +/- Cancer SMR has relevance to Cancer costs, but may be poorer at predicting future cancer costs than at explaining past cancer costs due to the strong connection between past costs and eventual cancer deaths |
| **Stability** |  |
| **Responsiveness** | + In principle, specifically-selected variables should perform better, but this is also subject to over-fitting |
| **Face Validity** | - The third strongest variable for Cancer (All-cause SMR) had counter-intuitive negative coefficients in the regression |

**6.3 Diagnostic groups vs Whole Acute**

To see whether separate diagnostic groups or the ‘Whole Acute’ option give better predictions, the separate diagnostic-group predictions must be aggregated to the Whole Acute level, to be compared with the predictions done using the ‘Whole Acute’ option. The aggregation should be done in the same way as in the NRAC formula: by summing the predicted ‘actual’ cost and the expected cost across the diagnostic groups, then dividing one by the other to produce an Acute predicted cost ratio.

The predictions can then be compared with the 2014/15 Whole Acute cost ratio, using RSS as before.

The results of this are shown in Table 14. There are no substantial differences in RSS between diagnostic groups and the Whole Acute option, for indices with 1, 2, or 3 indicators. For 4 indicators, it depends whether Whole Acute is predicted using the common index or the Whole Acute-specific index. Using the specific index, there is no difference, but using the common index, the RSS is a bit lower (i.e. the predictions are a bit better) for Whole Acute.

*Table 14: RSS obtained from comparing predictions derived from the top 1, 2, 3 and 4 variables from the restricted sets with the 2014/15 cost ratios – comparing Whole Acute with the aggregated predictions from separate diagnostic groups*

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Diagnostic groups – common index approach** | **Whole Acute – specific index approach** | **Whole Acute – common index approach** |
| RSS using top 1 index  | 355 | 355 | 355 |
| RSS using top 2 index | 409 | 408 | 408 |
| RSS using top 3 index | 355 | 354 | 354 |
| RSS using top 4 index | 408 | 408 | 344 |

Table 15 shows the coefficient of the needs index (i.e. the slope of the fitted line) across the diagnostic groups, for each index option. There are some consistent differences between the diagnostic groups, with Respiratory, Injury, Digestive and Heart having generally higher coefficients and Outpatients and Cancer having the lowest coefficient values. This shows that the strength of the relationship between the needs index and the cost ratios varies between diagnostic groups – although this does not mean that separate modelling improves the predictions (Table 14).

**Q: The Subgroup is asked to decide between retaining the diagnostic groups, or carrying out the Acute MLC adjustment at the whole care programme level.**

*Table 15: Coefficients of the needs index options*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Whole Acute** | **Cancer** | **Heart** | **Digestive** | **Injury** | **Other** | **Respiratory** | **Outpatients** |
| Coefficient of common top 4 index | 0.068 | 0.037 | 0.090 | 0.085 | 0.098 | 0.073 | 0.138 | 0.025(prison dummy 0.251) |
| Coefficient of specific top 4 index | 0.057 | 0.056 | 0.073 | 0.085 | 0.078 | 0.073 | 0.140 | 0.031(prison dummy 0.275) |
| Coefficient of common top 3 index | 0.078 | 0.042 | 0.099 | 0.100 | 0.107 | 0.086 | 0.156 | 0.032(prison dummy 0.243) |
| Coefficient of specific top 3 index | 0.078 | 0.056 | 0.099 | 0.103 | 0.107 | 0.086 | 0.163 | 0.035(prison dummy 0.257) |
| Coefficient of common top 2 index | 0.099 | 0.057 | 0.121 | 0.125 | 0.134 | 0.110 | 0.192 | 0.041(prison dummy 0.248) |
| Coefficient of specific top 2 index | 0.099 | 0.088 | 0.121 | 0.125 | 0.134 | 0.110 | 0.204 | 0.045(prison dummy 0.227) |
| Coefficient of common top 1 index | 0.172 | 0.084 | 0.222 | 0.224 | 0.231 | 0.188 | 0.348 | 0.076(prison dummy 0.201) |
| Coefficient of specific top 1 index | 0.172 | 0.140 | 0.222 | 0.224 | 0.231 | 0.188 | 0.367 | 0.076(prison dummy 0.201) |

**7. Next steps**

This work concludes the methodology outlined in Annex A. Once decisions are made to select (or narrow down) the index option(s), the next step will be to more fully evaluate the performance of the potential new model, including analysis across different urban-rural settings and age groupings. This will be presented at the May meeting.

**Annex A: Methodology for selecting the needs indicators**

The methodology adopted for selecting the needs indicators, based on the 2007 method from Technical Report D, is outlined below.

Preliminary selection of the candidate variables and regressions

1. Allocate the candidate variables to the appropriate topic/category.
2. Births and deaths
3. Health/morbidity
4. Unpaid care and elderly living alone
5. Deprivation
6. For each topic, compute the inter-correlation.
7. Retain candidate variables with low inter-correlations 🡪 Go to step 4.
8. Group pairs of candidate variables with very high inter-correlations to form subgroups. 🡪 Go to step 3.
9. For each subgroup, compute the correlation with cost ratios.
10. Eliminate near duplicates for each subgroup – retain candidate variables that have the highest correlation with cost ratios for most diagnostic groups.
11. New candidate variables list is formed.
12. Regress (stepwise) cost ratios against supply model with all the new candidate variables.
13. Display all the statistically significant candidate variables for each diagnostic group in a table.

Developing need indexes for the diagnostic groups cost ratios

1. Undertake a factor analysis – using principal components – of the new candidate variables. Selection is based on the extent to which each of the ‘raw’ variables appears to reflect the factors or dimensions that are generated by the factor analysis.
2. Choose the most effective variables.
3. The common restricted variable set is formed.
4. Use table from step 5 to check that most of them had appeared in the full stepwise regressions.
5. Regress cost ratios against supply model with the common restricted variable set.
6. Compare with the original equations – in terms of goodness of fit and the general specification test.
7. Compute the coefficients from regressing these variables against the diagnostic groups cost ratios.
8. Highlight variables with coefficients relatively high within the diagnostic groups.
9. Select variables that appear often across the diagnostic groups and have high coefficients.
10. Compute the sum of z-scores for those selected variables to form an index.

Developing specific indexes for the diagnostic groups cost ratios

1. Find the top 1, 2, 3 and 4 variables as the combinations yielding the highest R2, for each diagnostic group separately.

Figure B.1 shows the above methodology in the form of a flowchart. The numbers correspond to the numbers above.



*Figure A.1: Flowchart summary of index development methodology*

1. <http://www.gov.scot/Publications/2015/08/7995/downloads> [↑](#footnote-ref-1)
2. This was implemented in the analysis by using the Bayesian Information Criterion as the measure of goodness-of-fit. [↑](#footnote-ref-2)
3. Services covered are: inpatient, daycase, outpatient, A&E, and community prescribing. [↑](#footnote-ref-3)