**TAGRA ACUTE MLC SUBGROUP (By email)**

**Potential Indicators and Supply Variables**

**Background**

At the 10th meeting of the Subgroup in August 2015, the potential candidate variables and supply variables were discussed in the context of the Health Inequalities Impact Assessment discussion with equality leads.

The additional candidate variables from 2011 census data that were suggested in paper TAMLC34 (unpaid care, bad general health, low education level, mental health condition, and job seekers) were approved for inclusion, as it was felt that they had a plausible link to need for Acute services and that their geographical distribution could be expected to be stable over the 10-year interval between censuses. It was suggested that while unpaid care was likely to reflect need, the *absence* of care where it is needed may also impact on Acute service need. To attempt to capture this, it was agreed that a measure of the number of elderly people living alone could be included, based on census data. The inclusion of this variable is also consistent with the aim to include more variables of relevance to the elderly population.

The equality leads advised that the use of a percentage of non-white residents as a variable was problematic, since the relationships between ethnicity and healthcare needs are much more complex than this.

A query was raised in paper TAMLC34 around whether to exclude the supply variables that had been shown during the NRAC review to perform poorly. AST were asked to check whether the Subgroup had previously made any explicit decisions on the set of supply variables.

Finally, AST were also asked to check whether the Scottish Patients at Risk of Readmission and Admission (SPARRA) tool had previously been excluded as a potential candidate variable.

This paper presents investigations to explore these suggestions and presents the final potential candidate variables and supply variables. **It includes two issues for approval by 12th November. If we do not hear from you, we will assume that you approve these recommendations and we will proceed with the planned regression analyses accordingly.**

**1. Summary**

This paper addresses the feedback and suggestions from the last sub-group meeting, and presents the final potential candidate variables and supply variables table. Section 2 presents the final list of potential candidate variables. Section 3 explores the strengths and weaknesses of the different options for the supply model and a recommendation is made to the Subgroup to retain the current NRAC supply model, consisting of the variables IPACX and OPACX as well as the Health Board dummies. Finally, in section 4, the methodology for the selection of the needs indicators in the 2007 NRAC review is summarised as a basis for the methodology that will be used in the current review.

**2. Potential candidate variables**

The variables that will be tested as potential predictors of additional needs due to morbidity and life circumstances are shown in Table 1, along with their current status. Most variables have now been calculated at the 2011 data zones.

The number of elderly people living alone, based on census data, is now included, following the suggestion at the last Subgroup meeting.

Following the feedback from equality leads that the ethnicity variable needed modification, the Scottish Government ScotStat report “Which ethnic groups have the poorest health? An Analysis of Health Inequality and Ethnicity in Scotland”[[1]](#footnote-1) has been used as a guide. The report suggests that health outcomes for Gypsy/Traveller and Pakistani ethnic groups are worse on average, for both sexes, compared to the overall population, so counts of these populations are now suggested to be used.

SPARRA is not included because it was ruled out previously by the Subgroup. Paper TAMLC09 says, *“On investigating using the SPARRA data, it was concluded that this will not be a suitable candidate variable for the Acute MLC update as the data produces a score (known as a SPARRA score) for each patient. Many of the risk factors and variables used in SPARRA are already considered for the update or are highly correlated with them”*.

**Q1: The Subgroup is asked to approve the change from a single “non-white” ethnicity variable to the use of Gypsy/Traveller and Pakistani ethnic group populations.**

*Table 1. Potential candidate variables*

|  |  |  |
| --- | --- | --- |
| **Variable** | **Details** | **Status** |
| Low birth weight births – as a fraction of population | 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). | Completed |
| Death rate 0-74 all causes | 5 financial years’ GRO 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 | Completed |
| Death rate 0-74 Cancer | Completed |
| Death rate 0-74 CHD | Completed |
| Death rate 0-74 Stroke | Completed |
| 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 | Completed |
| All cause SMR 0-69 | Completed |
| All cause SMR 0-74 | Completed |
| Cancer SMR 0-64 | Completed |
| Cancer SMR 0-69 | Completed |
| Cancer SMR 0-74 | Completed |
| Heart Disease SMR 0-64 | Completed |
| Heart Disease SMR 0-69 | Completed |
| Heart Disease SMR 0-74 | Completed |
| Respiratory SMR 0-64 | Completed |
| Respiratory SMR 0-69 | Completed |
| Respiratory SMR 0-74 | Completed |
| Digestive System SMR 0-64 | Completed |
| Digestive System SMR 0-69 | Completed |
| Digestive System SMR 0-74 | Completed |
| External Causes SMR 0-64 | Completed |
| External Causes SMR 0-69 | Completed |
| External Causes SMR 0-74 | Completed |
| Other SMR 0-64 | Completed |
| Other SMR 0-69 | Completed |
| Other SMR 0-74 | Completed |
| High Resource Individual counts – as a fraction of population | 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). | Completed |
| 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. Included because of its link to deprivation. | Completed |
| Ratio of Did Not Attend counts 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. | Completed |
| Patients receiving Diabetes prescriptions (fraction of population) | 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 | Completed |
| Patients receiving Dementia prescriptions (fraction of population) | Completed |
| Patients receiving Respiratory prescriptions (fraction of population) | Completed |
| Long-term illness | Data from Census 2011 question 20 – standardised by age and sex using 2011 MYE population. | Awaiting data |
| Mental health condition | Data from Census 2011 question 20 – standardised by age and sex using 2011 MYE population. Included because of the link between mental health conditions and need for Acute services. | Completed |
| Limiting long-term illness | 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’). | Completed |
| Long-term sick and not seeking work | Data from Census 2011 questions 24-28 – standardised by age and sex using 2011 MYE population. | Completed |
| Elderly people living alone | Data from Census 2011 (counts of people living alone by the 20 NRAC age bands, so several cut-off ages could be used). Expressed as a fraction of 2011 MYE population. | Awaiting data |
| Unpaid care | 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). | Awaiting data |
| Bad general health | Data from Census 2011 question 19 – standardised by age and sex using 2011 MYE population. Uses number of respondents answering ‘Bad’ or ‘Very bad’. | Awaiting data |
| Low education level | Data from Census 2011 question 23 – standardised by age and sex using 2011 MYE population. Uses number of respondents with low education levels (various options). Included for its relation to deprivation. | Awaiting data |
| Job seekers | Data from Census 2011 question 25 – standardised by age and sex using 2011 MYE population (possibly economically active population only – try both). Included for its relation to deprivation. | Completed |
| Ethnic group populations (Pakistani, Gypsy/Traveller) | Data from Census 2011 question 15 – simple fraction of 2011 MYE population. | Awaiting data |

*Table 2. Possible supply variables for investigation*

|  |  |  |
| --- | --- | --- |
| **Variable** | **Description** | **Details** |
| GPCount | Number of GPs serving each data zone. | Calculated using headcount of GPs per practice as at 04/09/2015 from General Practitioner Contractor Database, and April 2015 CHI data to map GP practices to data zones. |
| GPSup, GPSup5, GPSup10 | A function of both the number of GPs serving each data zone and the distance to the practices. Distance estimated using 1km, 5km and 10km “intrazonal cost”, respectively (see Annex A for full details). | Calculated using headcount of GPs per practice as at 04/09/2015 from General Practitioner Contractor Database, April 2015 CHI data to map GP practices to data zones, 2011 data zone population centroids, and locations of practices. |
| IA1, OA1 | Size of nearest inpatient / outpatient facility. | Calculated using numbers of SMR01/00 episodes in 13/14, 2011 data zone population centroids, and locations of facilities. |
| 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. | Calculated using numbers of SMR01/00 episodes in 13/14, 2011 data zone population centroids, locations of facilities, and 13/14 SMR01/00 data to map facilities and their served data zones. |
| IPAC, OPAC | Similar to IPACX and OPACX, but with an attempt to correct the size of the facility for capacity. | Calculated using numbers of SMR01/00 episodes in 13/14, 2011 data zone population centroids, locations of facilities, 13/14 SMR01/00 data to map facilities and their served data zones, and 2013 MYE populations. |
| Health Board dummy variables | Represents the effect of the health board. | Calculated using current data zone—Health Board best-fit lookup. |

**3. Supply model exploration**

Table 2 presents the supply variables that were previously explored in the 2007 NRAC review. Annex A gives the exact formulation of each variable.

The IA and OA measures were not recommended in Technical Report D of the NRAC review since they tend to be biased by institutions with very small capacity. Also, in relation to IPAC and OPAC, Technical Report D states:

*“Though intuitively attractive, this method appears to overcompensate for population effects – especially in the Western Isles and Shetland. Although the hospitals are relatively small, the numbers in the catchment populations are also small, leading to values for (population corrected supply) that exceed those in central Glasgow. Because of our concern at the scale of the attraction constrained correction, only the two uncorrected measures (OPACX and IPACX) were used in the index development.”*

NRAC finally adopted only IPACX and OPACX (in addition to the Health Board dummies) as the supply model for the formula. This supply model is used across all care programmes. However, more recent analyses in the Greater Glasgow & Clyde NRAC investigation by Sarah Barry and Alex McConnachie have suggested that we should explore the use of all supply variables again.

Some analysis has been carried out on the different supply variables (see Annex B for the full results). Firstly, correlations were examined between the cost ratios and the supply variables. It was found that IPACX and OPACX are correlated most strongly with the cost ratios; several of the other supply variables showed counter-intuitive negative correlations. Secondly, regression was carried out using individual (or pairs of) supply variables, with and without the needs indicators; and using sensible combinations of supply variables, with and without the needs indicators.

Across these analyses, IPACX and OPACX consistently perform well compared to the other supply variables. A supply model consisting of IPACX, OPACX and GPSup (in addition to the Health Board dummies) also produces good results; GPSup has close to zero correlation with IPACX and OPACX, meaning that it can be considered to represent a separate aspect of supply and does not duplicate their effect. However, the original supply model – IPACX & OPACX – performs just as well in a full regression, and to retain it would keep the Acute model consistent with what is used in the other care programmes. The TAGRA core criteria of practicality and face validity would suggest the latter choice.

**Q2: The Subgroup is asked to approve the suggestion to retain the existing supply variables (IPACX and OPACX, plus health board dummies) for the Acute MLC model.**

**4. Methodology for selecting the needs indicators**

This section summarises the 2007 methodology for selecting the needs indicators, from Technical Report D. AST have discussed the methodology and regard it as a good basis for the Acute MLC review.

Preliminary selection of the candidate variables and regressions

1. Allocate the candidate variables to the appropriate topic/category.
2. Births and deaths
3. Income and Employment variables
4. Health/morbidity and providing unpaid care
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

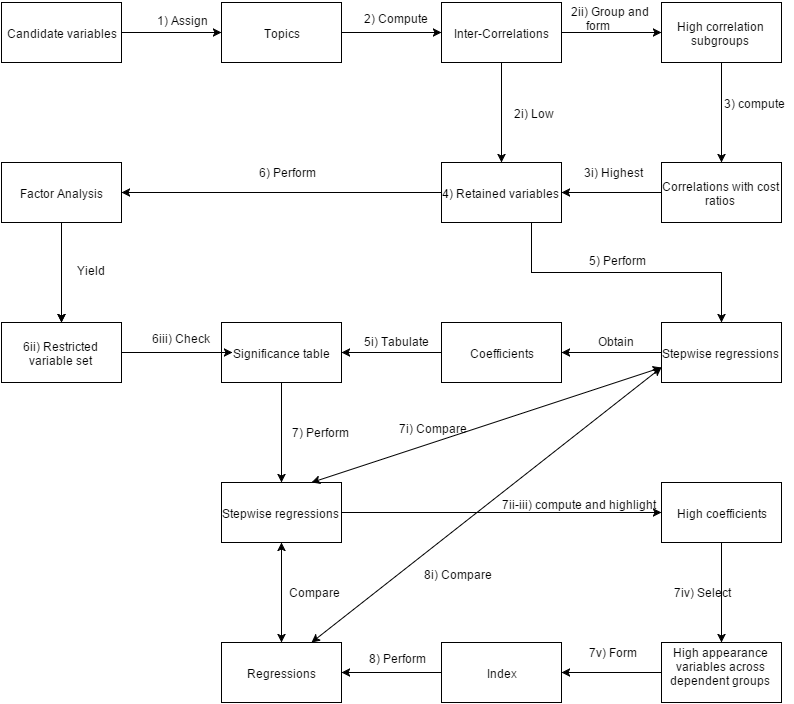
1. Undertake a factor analysis – using principal components – of the new candidate variables.

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.

1. Choose the most effective variables.
2. The restricted variable set is formed.
3. Use table from step 5 to check that most of them had appeared in the full stepwise regressions.
4. Regress (stepwise) cost ratios against supply model with all the restricted candidate variables.
5. Compare with the original equations – in terms of goodness of fit and the general specification test.
6. Compute the coefficients from regressing these variables against the diagnostic groups cost ratios.
7. Highlight variables with coefficients relatively high within the diagnostic groups.
8. Select variables that appear often across the diagnostic groups and have high coefficients.
9. Compute the sum of z-scores for those selected variables to form an index.
10. Regress cost ratios against supply model with the index.
11. Compare with the original equations – in terms of goodness of fit and the general specification test.

Figure 2 shows the above methodology in the form of a flowchart. The numbers correspond to the numbers above.

It is possible that the Subgroup may decide to allow the possibility of different needs indicators in different diagnostic groups. This was recently argued for in the context of the Health Inequalities Impact Assessment discussion. In this case, the final steps from 7)iii onwards would be revised.

**

*Figure 2: Flowchart summary of proposed index development methodology*

**Annex A: Supply variables calculation**

**A.1 Supply of/access to General Practice**

1. *“Number of GPs serving each Data Zone”.* This variable is the sum of the number of practices serving each DZ weighted by the whole time equivalent number of GPs in each practice.
2. *“GP supply, 1km intrazonal cost”, “GP supply, 5km intrazonal cost” and “GP supply, 10km intrazonal cost”.* 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.

After extensive enquiry, it seems that whole time equivalent GP numbers are no longer available by GP practice. We have therefore used GP headcount instead.

**A.2 Hospital supply**

1. *“Inpatient/Outpatient Access – nearest facility”.* Size of the 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 DZ population centroid.
2. *“IPACX/OPACX”.* A function of the linear distance from the population grid centroid of the DZ to each inpatient (I) or outpatient (O) facility it is served by. This can be expressed algebraically as:
3. *“IPAC/OPAC”.* The same inverse decay function is used as for IPAXC and OPACX, but the volume of each hospital is “attraction constrained”: the number of episodes at each hospital is weighted according to the size of the population it serves. This is the number of people in each DZ, weighted by the inverse square of the distance from the DZ centroid:

**A.3 Health Board dummy variables**

There is one binary variable representing each Health Board; it takes the value ‘1’ for data zones within the Health Board and ‘0’ otherwise.

**Annex B: Supply variables evaluation**

In this Annex we explore the performance of the candidate supply variables, and recommend, based on the results, that the existing supply model (IPACX and OPACX) should be retained.

**3.1 Correlations among supply variables, cost ratios and current acute index**

The correlations between supply variables and cost ratios for each diagnostic group are shown in Table 3. The correlations of ‘Outpatients’ with IPACX, OPACX or GPCount are moderately positive, but weakly negative for the correlation of ‘Outpatients’ with IPAC. The correlations of ‘Digestive’, ‘Other’, ‘Respiratory’ and ‘Current Acute Index’ each with IPACX or OPACX are weakly positive (highlighted in blue; negative correlations are highlighted in red); but none of the remaining correlations are high.

Only IPACX, OPACX, IA1, IA2 and GPCount do not have any counter-intuitive negative correlations with the cost ratios. IPACX and OPACX are correlated stronger to the cost ratios than the other supply variables. This suggests that IPACX and OPACX are capturing supply better than the other supply variables.

*Table 3. Correlations between supply variables and cost ratios.*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Correlations** | **Cancer cost ratios** | **Heart cost ratios** | **Digestive cost ratios** | **Injury  cost  ratios** | **Other cost ratios** | **Respiratory cost ratios** | **Outpatients cost ratios** | **Current Acute Index** |
| **IPACX** | 0.187 | 0.140 | 0.292 | 0.129 | 0.269 | 0.299 | 0.353 | 0.290 |
| **OPACX** | 0.167 | 0.118 | 0.221 | 0.121 | 0.212 | 0.277 | 0.401 | 0.247 |
| **IPAC** | -0.040 | -0.046 | -0.011 | -0.056 | 0.021 | -0.117 | -0.195 | -0.054 |
| **OPAC** | -0.058 | -0.098 | -0.074 | -0.097 | -0.063 | -0.150 | -0.141 | -0.113 |
| **IA1** | 0.076 | 0.071 | 0.170 | 0.033 | 0.104 | 0.126 | 0.090 | 0.144 |
| **OA1** | 0.019 | -0.026 | -0.021 | -0.013 | -0.009 | 0.009 | 0.033 | -0.008 |
| **IA2** | 0.087 | 0.075 | 0.161 | 0.046 | 0.088 | 0.137 | 0.134 | 0.130 |
| **OA2** | 0.029 | -0.017 | 0.007 | -0.009 | 0.014 | 0.019 | 0.048 | -0.024 |
| **GPCount** | 0.127 | 0.109 | 0.137 | 0.086 | 0.114 | 0.193 | 0.305 | 0.148 |
| **GPSup** | 0.034 | 0.064 | -0.015 | 0.087 | 0.019 | 0.057 | 0.043 | 0.084 |
| **GPsup5** | 0.015 | 0.043 | -0.041 | 0.056 | -0.030 | 0.030 | 0.011 | 0.029 |
| **GPSup10** | 0.004 | 0.030 | -0.055 | 0.038 | -0.054 | 0.012 | -0.011 | -0.003 |

Tables 4 and 5 show the inter-correlations among the hospital supply variables and the General Practice supply variables, respectively. As we would expect, there are strong associations among GPSup, GPSup5 and GPSup10 (highlighted in red). This is not the case for the hospital supply variables.

*Table 4. Correlations among hospital supply variables.*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Correlations** | **IPACX** | **OPACX** | **IPAC** | **OPAC** | **IA1** | **OA1** | **IA2** | **OA2** |
| **IPACX** | 1.000 | 0.961 | -0.385 | -0.428 | 0.500 | 0.245 | 0.460 | 0.343 |
| **OPACX** | 0.961 | 1.000 | -0.456 | -0.446 | 0.462 | 0.269 | 0.421 | 0.364 |
| **IPAC** | -0.385 | -0.456 | 1.000 | 0.736 | -0.171 | -0.116 | -0.219 | -0.140 |
| **OPAC** | -0.428 | -0.446 | 0.736 | 1.000 | -0.183 | -0.100 | -0.227 | -0.151 |
| **IA1** | 0.500 | 0.462 | -0.171 | -0.183 | 1.000 | 0.342 | 0.552 | 0.406 |
| **OA1** | 0.245 | 0.269 | -0.116 | -0.100 | 0.342 | 1.000 | 0.181 | 0.593 |
| **IA2** | 0.460 | 0.421 | -0.219 | -0.227 | 0.552 | 0.181 | 1.000 | 0.310 |
| **OA2** | 0.343 | 0.364 | -0.140 | -0.151 | 0.406 | 0.593 | 0.310 | 1.000 |

*Table 5. Correlations among General Practice supply variables.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Correlations** | **GPCount** | **GPSup** | **GPSup5** | **GPSup10** |
| **GPCount** | 1.000 | 0.096 | 0.142 | 0.154 |
| **GPSup** | 0.096 | 1.000 | 0.890 | 0.813 |
| **GPSup5** | 0.142 | 0.890 | 1.000 | 0.985 |
| **GPSup10** | 0.154 | 0.813 | 0.985 | 1.000 |

Table 6 shows the correlations between GP and hospital supply variables. We can see GPCount is quite highly correlated with IPACX and OPACX. GPSup, GPSup5 and GPSup10, on the other hand, all have close to zero correlation with IPACX and OPACX. This means that they might explain supply variations other than the supply variations explained by IPACX and OPACX.

*Table 6. Correlations between GP and hospital supply variables.*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Correlations** | **IPACX** | **OPACX** | **IPAC** | **OPAC** | **IA1** | **OA1** | **IA2** | **OA2** |
| **GPCount** | 0.741 | 0.782 | -0.519 | -0.547 | 0.301 | 0.173 | 0.348 | 0.251 |
| **GPSup** | 0.043 | 0.111 | -0.210 | -0.217 | -0.003 | 0.014 | -0.028 | 0.004 |
| **GPSup5** | 0.046 | 0.119 | -0.297 | -0.298 | 0.018 | 0.033 | 0.015 | 0.034 |
| **GPSup10** | 0.032 | 0.104 | -0.318 | -0.319 | 0.020 | 0.035 | 0.025 | 0.041 |

**3.2 Performance of supply variables in regression analysis**

Adjusted R-squared values are used as a goodness of fit measure, indicating the amount of variation in the observations that is explained by a model. The performance of the different supply variables (or pairs of variables) when entered alone into the regression model is shown in Table 7 below. The first half of the table represents hospital supply variables (highlighted with a light green cell colour) and the second half represents supply of General Practice (highlighted with a light orange cell colour). The highest adjusted R-squared values for each diagnostic group in each supply category are highlighted in red.

*Table 7. Supply variables performances.*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Adjusted**  **R-squared** | Predictors | Cancer | Heart | Digestive | Injury | Other | Respiratory | Outpatients |
| **3 years Data Zones 2011** | supply model (IA2 & OA2) | 6.8% | 2.7% | 17.9% | 3.5% | 14.9% | 9.7% | 40.4% |
| supply model (IA1 & OA1) | 6.6% | 2.5% | 17.6% | 3.3% | 14.8% | 9.2% | 39.8% |
| supply model (IPAC & OPAC) | 6.7% | 2.7% | 17.7% | 3.4% | 15.4% | 10.4% | 39.5% |
| supply model (IPACX & OPACX) | 6.9% | 3.2% | 18.2% | 4.0% | 15.9% | 10.9% | 40.3% |
| supply model (GPCount) | 6.8% | 2.5% | 17.4% | 3.0% | 14.6% | 9.2% | 40.3% |
| supply model (GPSup) | 6.8% | 2.8% | 17.7% | 3.4% | 15.2% | 9.3% | 39.3% |
| supply model (GPSup5) | 6.7% | 2.5% | 17.5% | 3.1% | 14.8% | 9.1% | 39.3% |
| supply model (GPSup10) | 6.6% | 2.3% | 17.5% | 3.0% | 14.6% | 8.9% | 39.3% |

It is very clear that out of all the different supply models from both supply categories, the supply model with IPACX and OPACX produces the highest adjusted R-squared values across all diagnostic groups. This coincides with our initial impression and we can conclude that IPACX and OPACX are probably still the most appropriate supply variables as the basis of a supply model.

For the supply of General Practice category, “GP supply, 1km intrazonal cost” has the highest adjusted R-squared values among all diagnostic groups, except for ‘Outpatients’. Hence, it may be worth looking into a supply model with the combination of IPACX, OPACX and GPSup. This is further investigated in sections 3.4 and 3.5.

**3.3 Performance of supply variables together with needs indicators**

The performance of the different supply variables (or pairs of variables) when entered along with the current needs indicators into the regression model has also been examined. Each supply option was combined in turn with (i) SMR only, (ii) LLTI only, (iii) SMR + LLTI, and (iv) the current needs index, giving four groups of models.

In general, it appears that models with either IPACX & OPACX or IPAC & OPAC perform better than models with the other supply variables; and there is little difference between these two options, when combined with the indicators. However, Table 3 showed that IPAC & OPAC are negatively correlated with almost all of the diagnostic groups’ cost ratios, which is counter-intuitive. We therefore conclude that IPACX and OPACX are the most appropriate overall supply variables. (Note also the limitation of IPAC & OPAC mentioned in Technical Report D.)

**3.4 Performance of combined supply models**

In section 3.2 and section 3.3, we have found that the supply model with IPACX and OPACX performs the ‘best’. Since IPACX and OPACX represent the hospital supply, it would not make sense to combine them with any other pairs of hospital supply variables. However, supply of General Practice represents another supply element. It would be sensible to choose the most effective General Practice supply variable; and combine it with IPACX and OPACX. In simple terms, we want the ‘best’ pair of supply variable from hospital supply variables and the ‘best’ supply variable from supply of General Practice.

From the correlations in Table 6, we can regard GPCount as inappropriate for combination with IPACX and OPACX. We could, however, combine IPACX, OPACX and (GPSup or GPSup5 or GPSup10) to form a supply model. The performance of these combined supply models when entered alone into the regression model are displayed in Table 8 below. The highest adjusted R-squared values for each diagnostic group are highlighted in red.

*Table 8. Supply models performances with different supply variables combinations.*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Adjusted**  **R-squared** | Predictors | Cancer | Heart | Digestive | Injury | Other | Respiratory | Outpatients |
| **3 years Data Zones 2011** | supply model (IPACX & OPACX) | 6.9% | 3.2% | 18.2% | 4.0% | 15.9% | 10.9% | 40.3% |
| supply model (IPACX & OPACX + GPSup) | 7.0% | 3.7% | 18.5% | 4.3% | 16.5% | 11.3% | 40.3% |
| supply model (IPACX & OPACX + GPSup5) | 6.9% | 3.4% | 18.3% | 4.0% | 16.0% | 11.1% | 40.3% |
| supply model (IPACX & OPACX + GPSup10) | 6.9% | 3.3% | 18.2% | 3.9% | 15.9% | 11.0% | 40.4% |
| supply model (IA1 & OA1) | 6.6% | 2.5% | 17.6% | 3.3% | 14.8% | 9.2% | 39.8% |
| supply model (IA1 & OA1+ GPSup) | 6.8% | 3.1% | 17.9% | 3.6% | 15.3% | 9.7% | 39.9% |
| supply model (IA1 & OA1+ GPSup5) | 6.6% | 2.8% | 17.7% | 3.3% | 14.9% | 9.4% | 39.8% |
| supply model (IA1 & OA1+ GPSup10) | 6.6% | 2.7% | 17.7% | 3.3% | 14.8% | 9.3% | 39.8% |
| supply model (IA2 & OA2) | 6.8% | 2.7% | 17.9% | 3.5% | 14.9% | 9.7% | 40.4% |
| supply model (IA2 & OA2+ GPSup) | 7.0% | 3.3% | 18.2% | 3.9% | 15.5% | 10.2% | 40.4% |
| supply model (IA2 & OA2+ GPSup5) | 6.8% | 3.0% | 18.0% | 3.6% | 15.1% | 9.9% | 40.4% |
| supply model (IA2 & OA2+ GPSup10) | 6.8% | 2.8% | 17.9% | 3.5% | 15.0% | 9.8% | 40.4% |

From Table 8, a combined supply model with IPACX, OPACX and GPSup produces the highest adjusted R-squared values across all diagnostic groups.

**3.5 Performance of combined supply model together with needs indicators**

Table 9 compares the effect of using (IPACX and OPACX) vs (IPACX, OPACX and GPSup) as the supply model in a combined regression with the current needs indicators. The highest adjusted R-squared values across all models are highlighted in red.

*Table 9. Index models performances with different supply models as basis.*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Adjusted**  **R-squared** | Predictors | Cancer | Heart | Digestive | Injury | Other | Respiratory | Outpatients |
| **3 years Data Zones 2011** | supply model (IPACX & OPACX) + SMR75 | 11.1% | 15.9% | 32.9% | 19.8% | 36.4% | 31.2% | 45.5% |
| supply model (IPACX & OPACX + GPSup) + SMR75 | 11.1% | 15.9% | 32.9% | 19.8% | 36.4% | 31.2% | 45.5% |
| supply model (IPACX & OPACX) + LLTI | 9.5% | 20.2% | 38.2% | 25.6% | 45.1% | 36.1% | 50.2% |
| supply model (IPACX & OPACX + GPSup) + LLTI | 9.5% | 20.2% | 38.2% | 25.6% | 45.1% | 36.1% | 50.3% |
| supply model (IPACX & OPACX) + SMR + LLTI | 11.1% | 21.1% | 39.1% | 26.6% | 46.1% | 38.0% | 50.2% |
| supply model (IPACX & OPACX + GPSup) + SMR + LLTI | 11.1% | 21.1% | 39.1% | 26.6% | 46.1% | 38.0% | 50.3% |
| supply model (IPACX & OPACX) + current acute index | 10.8% | 20.5% | 38.4% | 25.8% | 44.8% | 37.5% | 48.9% |
| supply model (IPACX & OPACX + GPSup) + current acute index | 10.8% | 20.5% | 38.4% | 25.8% | 44.8% | 37.5% | 49.0% |

Although it was clear from Table 8 that the combined supply model (IPACX, OPACX and GPSup) explains more of the cost ratios’ variation than the existing supply model (IPACX and OPACX only), when these are combined with the current indicators as in Table 9, there are almost no differences in the adjusted R-squared value.

The original supply model – IPACX & OPACX – is therefore preferable for the sake of simplicity and to be consistent with what is used in the other care programmes.

1. Available from <http://www.gov.scot/Publications/2015/08/7995/downloads> [↑](#footnote-ref-1)