**TAGRA ACUTE MLC SUBGROUP Tuesday 9th June 2015**

**UNMET NEED ANALYSIS**

**Background**

The NHSScotland Resource Allocation Committee (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 the existence and extent of any socio-economic inequities in healthcare utilisation and, where appropriate, to adjust the Formula to reflect such unmet need.

There is an unmet need adjustment incorporated in the NRAC Formula only for the Heart acute diagnostic group. To introduce the correction, the 25% most deprived Intermediate Zones are excluded when fitting the reference model.

Unmet need was discussed most recently at the March Acute MLC subgroup meeting (TAMLC 24). The group agreed that AST should replicate the two-stage shortfall method analysis reported in McConnachie and Sutton (2004)[[1]](#footnote-1) with up-to-date Scottish Health Survey (SHeS) data. The SHeS diagnoses data allows for a match to four out of the seven acute diagnostic groups; therefore the unmet need analysis has been performed for Cancer, Digestive, Heart and Respiratory. The analysis has been carried out at Intermediate Zone level because the SHeS data contains many Data Zones without a positive response to the question of a long-term condition, which poses a problem for modelling.

**1. Summary**

This paper presents the results of testing for unmet need four acute diagnostic groups – Cancer, Digestive, Heart and Respiratory. Section 2, presents information on the relationship between healthcare utilisation (cost ratios) and the current acute index of need (a combination of SMR in ages 0-74 and LLTI) and sets out the NRAC reference model results. Section 3 describes the results of linear regressions of cost ratios upon the acute index along with upper and lower splines for the acute index (known as simple shortfall method which was proposed in the first Arbuthnott Report in 1999). Section 4 presents the result of the two-stage shortfall method, i.e. the SHeS data is used to calculate predicted prevalence and then the cost ratios are regressed on the predicted prevalence along with upper and lower splines for the acute index. Section 5 discusses the possible next steps for the unmet need analysis and provides some questions for the Subgroup to consider.

**2. NRAC Reference model**

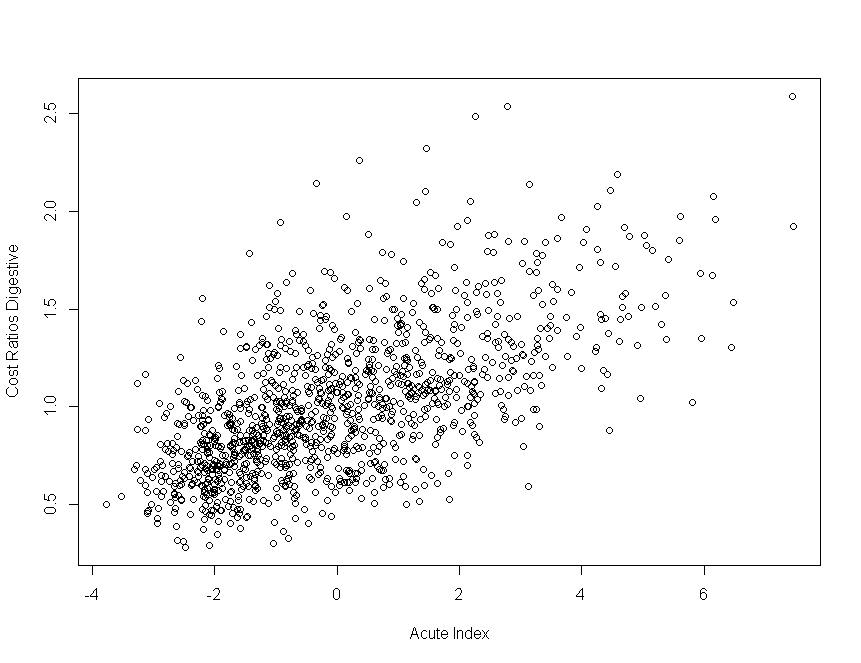
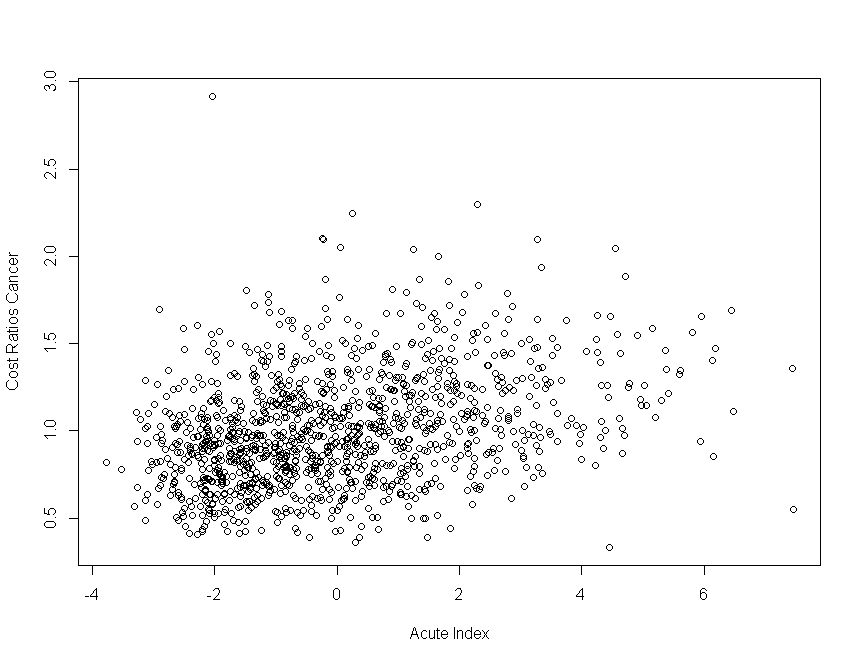
The adjustments for Morbidity and Life Circumstances in the formula are derived using regression models that include Health Board dummy variables, supply variables and the acute index (the sum of the z-scores of the all cause Standardised Mortality Ratio (SMR) in ages 0-74 and the Limiting Long-Term Illness (LLTI) ratio). This set of variables is referred to as the reference model in this work.

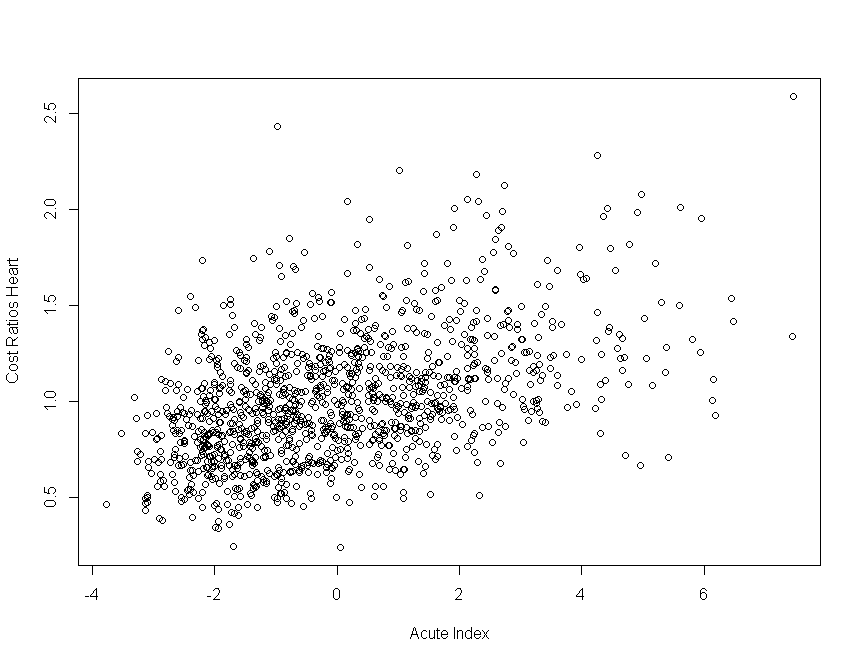
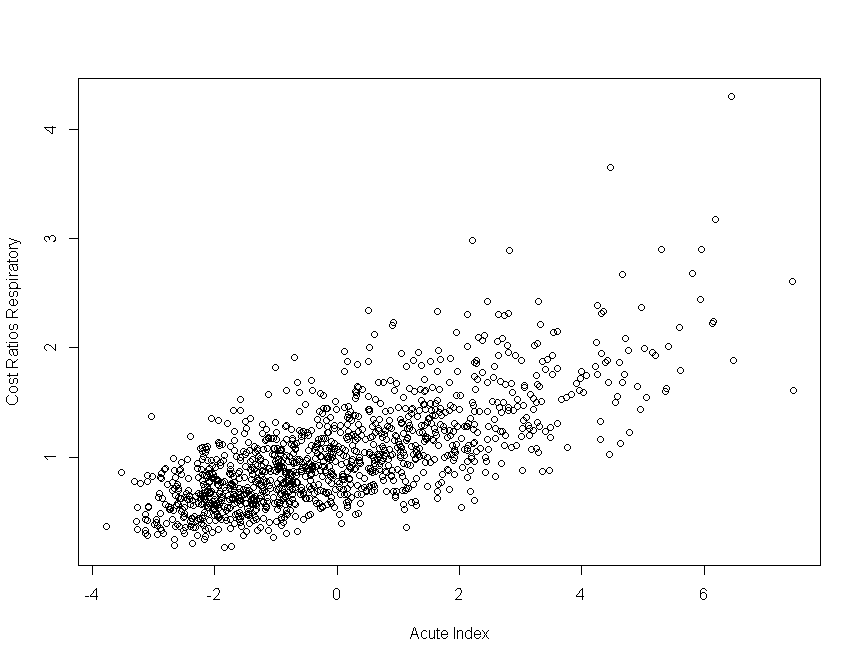
The relationship between healthcare utilisation (cost ratios) and the reference variables is assumed to be linear and the model is expressed algebraically as

***Cost Ratios = Health Board dummies + Supply + Acute Index + errors***

in which the healthcare utilisation is represented by the ratio of the actual costs of healthcare (taking into account activity type and the length of stay in that specific neighbourhood) to the expected costs (based on the neighbourhood’s population and the national average costs per head).

The cost ratios are calculated at Intermediate Zones for 2011/12 financial year. The supply variables are calculated in 2009 using 2001 Census data. The needs indicators included in the acute index are both age-sex adjusted ratios; the SMR is calculated using death records from 2008 to 2012 calendar years, and the LLTI ratio is calculated using 2011 Census data.

The relationship between the cost ratios and the acute index at Intermediate Zone level for the four diagnostic groups of interest (Cancer, Digestive, Heart and Respiratory) are plotted and applied below (Figure 2.1).



*Figure 2.1 Scatter plots of cost ratios against the acute index by diagnostic group.*

The scatter plots do not imply any clear deviations from the linear relationship for Digestive and Heart. The scatter plots suggest that there might be no further effect in the cost ratios for acute index values greater than around 3 for Cancer and a sharper increase in the cost ratios for acute index values greater than around 4 for Respiratory. However, for all four diagnostic groups the data points are very sparse at the extremes and the scatter plots do not clearly suggest any unmet/overmet need effects. These have been tested formally and the results are included in Sections 3 and 4 of the current paper.

The reference model performs well for all four diagnostic groups and the acute index is significant in all cases (p-values less than 0.05). The adjusted R2 values – the percentage of variance in the cost ratios that is explained by the model – are 23.45%, 55.02%, 27.00% and 55.59% respectively for Cancer, Digestive, Heart and Respiratory.

**3. Simple shortfall method**

The simple shortfall method – proposed in the first Arbuthnott Report in 1999 – is based on the linearity assumption. That is, it assumes that there should be a linear relationship between the cost ratio and the indicators of need. It effectively tests whether that relationship is in fact constant, across the full range of need, in the data or whether there is a significant change in the slope of healthcare use (cost ratios) at high values of the acute index. In this paper, the simple shortfall method is used to test for significant change in the slope at both upper and lower ends of the range of the acute index (the cut-off points are restricted to lie between 5% and 30% at either end). So, the model fitted is an extension of the reference model, allowing for shortfall through the inclusion of additional spline terms (i.e. allowing for different slopes in the regression line) for areas with the lowest and the highest acute index values:

***Cost ratios = HB dummies + Supply + Acute Index + Acute Index splines + errors.***

Additional details on the calculation of the spline terms are included in Annex A. Regressions are fitted for all possible combinations of cut-off points and the ‘best’ cut-off points are chosen so that the model has the highest explanatory power, i.e. the highest adjusted R2 value. The detailed results of fitting the simple shortfall method are included in Annex B as well as a comparison with the reference model.

In summary, the simple shortfall method shows evidence for **significant unmet need effects for Cancer** in the 15% IZs with the highest acute index values and **significant overmet need effects for Respiratory** in the 5% IZs with the highest acute index values. There is no evidence of unmet/overmet need in healthcare use for Digestive and Heart diagnostic groups.

A possible correction for unmet/overmet need – discussed within AST – is to extrapolate a regression line based on the areas not thought to be affected to all areas. For Cancer, if the reference model is fitted - excluding the 15% IZs with the highest acute index values – the regression coefficient for the acute index is increasing from 0.038 to 0.050, reflecting the adjustment for unmet need in areas with the highest acute index values. Respectively for Respiratory – excluding the 5% IZ with the highest acute index values – the regression coefficient for the acute index is decreasing from 0.163 to 0.157, reflecting the adjustment for overmet need in areas with the highest acute index values. The extent of the adjustments to the overall slope is relatively modest, especially for Respiratory. However, the resource implications cannot be checked at this stage of the Review.

**4. Two-step shortfall method**

The two-step shortfall method – proposed in the McConnachie and Sutton (2004) paper – involves a two stage process. The first step is to regress the SHeS morbidity prevalence data (by diagnostic group) on indicators of need (i.e. the acute index) and thereby calculate predicted morbidity prevalence for all Intermediate Zones. In the second step the utilisation data (cost ratios) are regressed on the predicted prevalence along with upper and lower splines for the acute index. The spline parameters can then be used to test for under or over utilisation at both ends of the acute index spectrum.

**4.1. Scottish Health Survey data**

The Scottish Health Survey data was obtained, and some preliminary analysis was undertaken and summarised in paper TAMLC 24. The SHeS data used in the analysis is from 2008 to 2011 surveys and contains around 37,000 respondents. The data are coded by ICD10 which allows a match to four acute diagnostic groups: Cancer, Digestive, Heart and Respiratory. Positive responses to the survey question are highest for Heart, for which around 10% of responses are positive. Only four Intermediate Zones have no respondents at all.

*Table 4.1 Summary of responses to SHeS 2008-2011 at Intermediate Zones by diagnostic group.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Diagnostic  group | Mean number  of respondents | Number of IZs  with no positive responses | Mean number of respondents for IZs  with no positive responses | Mean number of respondents for IZs with positive responses |
| Respiratory | 29.97 | 214 (17%) | 23.13 | 31.41 |
| Heart | 29.97 | 160 (13%) | 21.56 | 31.22 |
| Cancer | 29.97 | 774 (63%) | 26.15 | 36.45 |
| Digestive | 29.97 | 422 (34%) | 24.48 | 32.84 |

The mean number of respondents for Intermediate Zones with no positive responses (zero estimated prevalence) is lower than the overall number of respondents (29.97) for all diagnostic groups. This can be considered an issue of bias due to different numbers of respondents in different IZs. It could be accounted for by using a weighted regression, but this has not been explored because of the complexity of the modelling and the time constraint.

The distribution of the number of respondents across the acute index values is compared to the population distribution and the results are summarised in Table 4.2. The total population used is the average of the 2008 to 2011 mid-year population estimates.

*Table 4.2. Comparison of percentage of Scottish population and percentage of SHeS respondents across acute index values.*

|  |  |  |  |
| --- | --- | --- | --- |
| Acute Index | Number of IZs | Percentage  of total population | Percentage  of SHeS respondents |
| (-4; -2] | 186 | 15.37 | 14.82 |
| (-2; 0] | 500 | 40.22 | 43.67 |
| (0; 2] | 351 | 28.66 | 26.48 |
| (2; 4] | 145 | 11.98 | 11.28 |
| (4; 6] | 42 | 3.25 | 3.27 |
| (6; 8] | 7 | 0.52 | 0.47 |

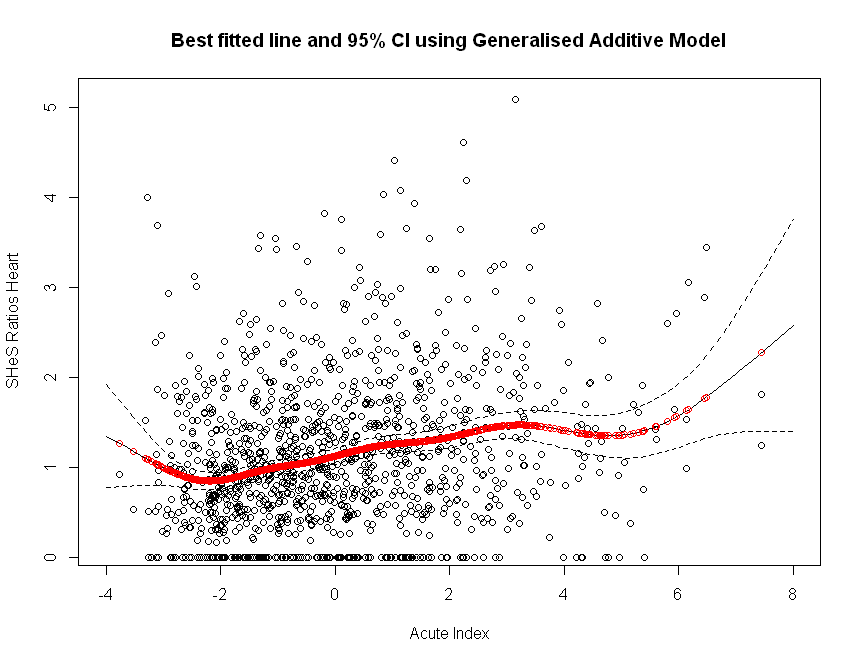
The distribution of the respondents across acute index values is similar to the Scottish population distribution. This implies that there is no strong bias in terms of the survey coverage. However, it is worth mentioning that there might be self-reporting bias which may vary as a function of the acute index.

The SHeS data is used to calculate age-sex adjusted rates of morbidity (same methodology used as for the calculation of the acute index constituent variables) and this is used as an estimate of prevalence in the analysis.

**4.2. Predicted Prevalence and the two-step shortfall method**

The first step uses the SHeS morbidity data as the dependent variable and the acute index as an independent variable and fits a non-linear model predicting morbidity prevalence rates.

It was found that the choice and the type of the non-linear model and its complexity strongly affect the best fitted values. Furthermore, the large number of zeros SHeS data points affects the fit as well, especially at the lowest and highest values of the acute index. As one example of a non-linear model, Generalised Additive Models are fitted for all four diagnostic groups. The acute diagnostic group Heart is chosen as an example and the scatter plot below summarises the results of the non-linear modelling (Figure 4.1). Scatter plots for Cancer, Digestive and Respiratory are included in Annex C.

****

*Figure 4.1. Scatter plot of SHeS ratios against the acute index for the acute diagnostic group Heart. Predicted values from fitting a Generalised Additive Model shown in red. Dashed lines represent 95%confidence intervals.*

The deviations from the line at both low and high values of the acute index are caused by only a few data points. The data is very sparse at both ends of the acute index and the best fitted line for different non-linear models (e.g. cubic splines regression or generalised additive models with different degrees of freedom) give substantially different results. This is the case for all diagnostic groups and the choice of the non-linear model is seen to result in different cut-off points for the spline variables in the second stage of the method.

The second stage in the two-step shortfall method can be expressed algebraically as

***Cost ratios = HB dummies + Supply + Predicted Prevalence***

***+ Acute Index splines + errors.***

The detailed results of fitting the two-step shortfall method are included in Annex D. Given the quality of the SHeS data and the fact that the predicted prevalence values are quite “method-sensitive”, the results from the two-step shortfall method might not be robust. The outcomes from fitting the simple and the two-step shortfall methods are similar for Respiratory and Cancer (significant unmet need for Cancer and significant overmet need for Respiratory in the deprived end). For Digestive and Heart, the outcomes are different and the two-step shortfall method identifies significant effects at both low and high values of the acute index; however, the outcomes are in the opposite direction to expectations – significant unmet need in areas with the lowest acute index and overmet need in areas with the highest acute index. What is happening here is that the spline variables are re-adjusting for the patterns imposed from the acute index – prevalence relationships, in which prevalence is seen to increase at the lowest acute index values (e.g. Figure 4.1).

**4.3. Comparison of results**

The two-step shortfall method has been explored in the current work and the results are presented in Section 4.2. However, it is worth noting that although the method is based on McConnachie and Sutton (2004), there are some significant differences between the current implementation and the one used in McConnachie and Sutton (2004):

* The predictors of prevalence are different. The current acute index contains all cause SMR in ages 0-74 and LLTI. However, in 2004, the Arbuthnott index was usedand it contained not only morbidity variables:
  + all cause standardised mortality ratios in ages 0-64,
  + the proportion of households having 2 or more deprivation indicators,
  + the proportion of the population of working age claiming unemployment benefit,
  + the proportion of the population ages 65 and over claiming income support;
* Arbuthnott areas were used for the analysis in 2004 (717 Arbuthnott areas) and Intermediate Zones have been used in the current work (1235 IZs); Also, respondents were sampled from 451 out of 717 Arbuthnott areas and from 1231 out of 1235 Intermediate zones.
* Data from 1995 and 1998 waves of the SHeS were used by McConnachie and Sutton while in the current analysis four years of SHeS data are used (2008-2011).

These differences could lead to variations in the results from the non-linear modelling since the predictors of the prevalence (Arbuthnott index and acute index) are not equivalent and there are more zero observations at the smaller geography unit.

In 2004, McConnachie and Sutton found significant unmet need effects using the two-step shortfall method for Heart, Cancer and Respiratory and significant overmet need effects for Digestive in the most deprived areas. They also found significant unmet need effects for Cancer and Respiratory in the most affluent areas. These results are not generally consistent with those obtained for the present analysis, which could reflect (i) methodological sensitivity in the two-step shortfall method, (ii) the difference in acute index, geographical scale and level of time-averaging, and / or (iii) real changes through time in the utilization – acute index and/or prevalence - acute index relationships. In 2004, McConnachie and Sutton only applied the simple shortfall method to the Heart diagnostic group and this did now show any significant unmet/overmet need effects. This result is consistent with the present analysis.

**5. Discussion**

The results of the analysis have been discussed within AST and with Professor Matt Sutton. It was agreed that the analysis presented in Section 4 provides evidence that the two-step shortfall method should not be used to identify unmet/overmet need across the acute diagnostic groups. The main reasons for this are (i) the current acute index contains different variables compared to the Arbuthnott index[[2]](#footnote-2), (ii) the large number of zero SHeS data points (zero prevalence), even at Intermediate Zone geographies, (iii) the choice of the type of the non-linear model affects the final cut-off points which introduces methodological subjectivity.

The simple shortfall method may in fact be considered to be the more reliable approach for the unmet need investigation. The presence of significant overmet need effects in areas with the highest acute index values for Respiratory might be due to severity issues, i.e. more expensive cases, but not more cases and not better access to services. The significant unmet need effects for Cancer are potentially more difficult to rationalise. Clinical input will be requested from ISD Consultants in Public Health for both Cancer and Respiratory unmet/overmet need before decisions are made whether to adjust for these effects.

The suggested next steps for the unmet need analysis are as follows

* Ask for clinical input from ISD Consultants in Public Health on the Cancer and Respiratory results
* Fit the simple shortfall method at Data Zones since this is the geography chosen for the Review and check the consistency of the results between Intermediate Zones and Data Zones
* Fit the simple shortfall method at redrawn Data Zones once the ‘new’ acute index is selected and check the validity of the results;

**The subgroup is asked to discuss the recommendation to use the simple shortfall method as the main method for the unmet need investigation and to agree the next steps for the analysis.**

**ANNEX A: Definition of spline variables**

The spline variables are included in the simple and two-step shortfall methods and they are defined using the acute index values (xi, where i is between 1 and 1235 for each Intermediate Zone). Splines are included for both most affluent (5th to 30th percentile points) and most deprived (70th to 95th percentile points) and they are noted respectively as kL and kH.

The spline for low acute index values (xiL) is defined as:

xiL = kL – xi if kL > xi

=0 if kL ≤ xi

And the spline for high acute index values (xiH) is similarly defined as:

xiH = xi – kH if xi > kH

=0 if xi ≤ kH

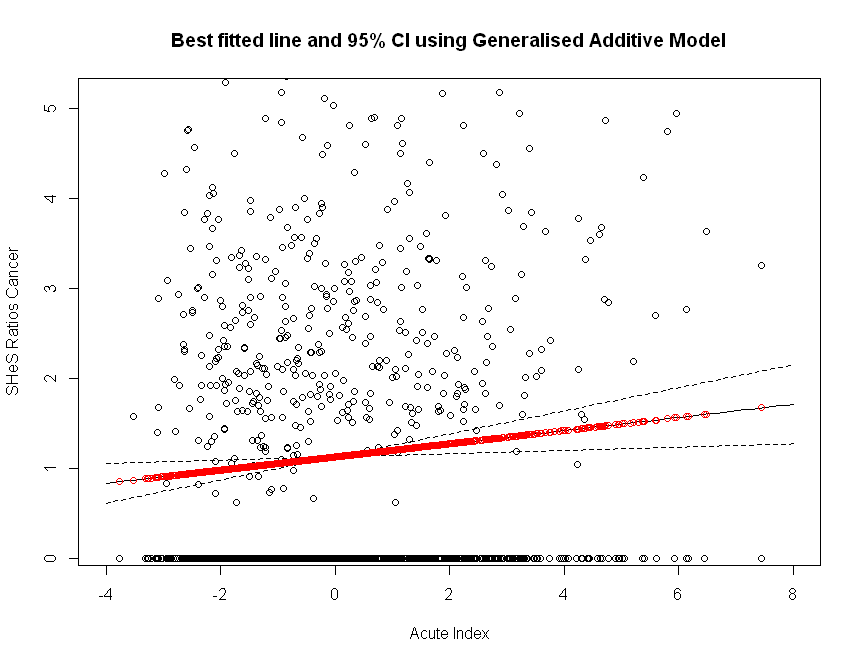
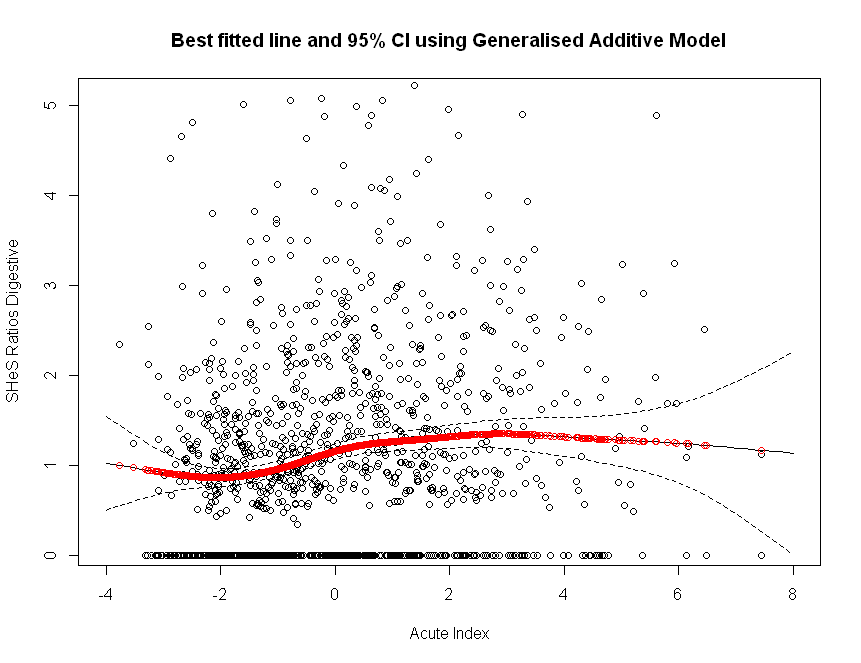
These spline variables are included in the linear regression model in addition to the supply variables, Health Board dummies and the acute index (simple shortfall method) or the predicted prevalence (two-step shortfall method). If the regression coefficients of both spline variables are not significantly different from 0, then there is no evidence of unmet need in health care use. If the sign of the regression coefficient of the spline variable (low or high acute index values) is negative and it is significant (p-value less than 0.05), then there is significant effect of unmet need. Otherwise, if the sign of the regression coefficient of the spline variable (low or high acute index values) is positive and it is significant (p-value less than 0.05), then there is significant effect of overmet need.

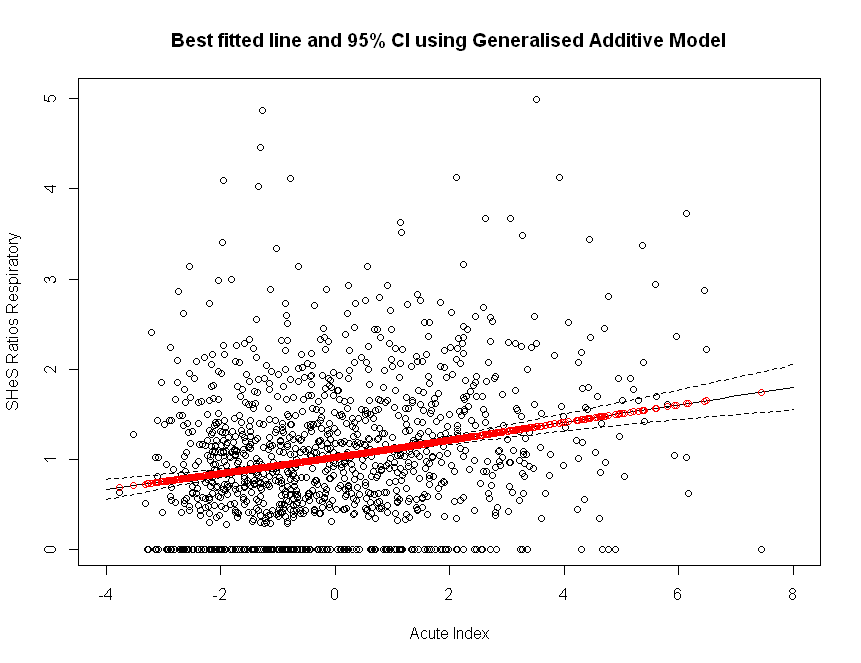
**ANNEX B: Simple shortfall method results table**

*Table B.1. Summary of results for the reference model (Model 1) and the simple shortfall method (Model 2) for Heart, Respiratory, Digestive and Cancer.*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Diagnostic Group** | **Model** | **Main effect**  **(Acute Index)** | **%**  **lower cut** | **Effect** | **%**  **upper cut** | **Effect** | **Adjusted R2** |
| Heart | 1 | 0.09  (0.08, 0.10)  p<0.0001 |  | | | | 27.00% |
| 2 | 0.09  (0.08; 0.10)  p<0.0001 | 5% | -0.10  (-0.30; 0.10)  p=0.31  **Not significant** | 90% | -0.03  (-0.07; 0.10)  p=0.21  **Not significant** | 27.07% |
| Respiratory | 1 | 0.16  (0.15; 1.73)  p<0.0001 |  | | | | 55.59% |
| 2 | 0.15  (0.14; 0.17)  p<0.0001 | 30% | -0.02  (-0.08; 0.04)  p=0.47  **Not significant** | 95% | 0.09  (0.02; 0.15)  p=0.007  **Overmet need** | 55.78% |
| Digestive | 1 | 0.10  (0.09; 0.11)  p<0.0001 |  | | | | 55.02% |
| 2 | 0.09  (0.08; 0.10)  p<0.0001 | 30% | -0.03  (-0.08; 0.01)  p=0.15  **Not significant** | 93% | 0.03  (-0.02; 0.07)  p=0.24  **Not significant** | 55.04% |
| Cancer | 1 | 0.04  (0.03; 0.05)  p<0.0001 |  | | | | 23.45% |
| 2 | 0.05  (0.03; 0.06)  p<0.0001 | 30% | -0.02  (-0.07; 0.06)  p=0.56  **Not significant** | 85% | -0.04  (-0.08;-0.01)  p=0.029  **Unmet need** | 23.85% |

**Annex C: Scatter plots of step 1 results of the two-step shortfall method**

****

****

*Figure C.1. Scatter plot of SHeS ratios against the acute index for the acute diagnostic groups Cancer, Digestive and Respiratory. Predicted values from fitting a Generalised Additive Model shown in red. Dashed lines represent 95% confidence intervals.*

**ANNEX D: Two-step shortfall method results table**

*Table C.1. Summary of results for the two-step shortfall method for Heart, Respiratory, Digestive and Cancer.*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Diagnostic Group** | **Main effect**  **(Acute Index)** | **%**  **lower cut** | **Effect** | **%**  **upper cut** | **Effect** | **Adjusted R2** |
| **Heart** | 0.62  (0.50; 0.74)  p<0.001 | 9% | -0.28  (-0.39;-0.18)  p<0.0001  **Unmet need** | 73% | 0.04  (0.02; 0.07)  p<0.001  **Overmet need** | 27.01% |
| **Respiratory** | 1.60  (1.44; 1.78)  p<0.0001 | 30% | -0.04  (-0.09; 0.02)  p=0.21  **Not significant** | 95% | 0.08  (0.01; 0.15)  p=0.02  **Overmet need** | 55.75% |
| **Digestive** | 0.52  (0.41; 0.63)  p<0.0001 | 30% | -0.12  (-0.15;-0.08)  p<0.001  **Unmet need** | 76% | 0.11  (0.09; 0.13)  p<0.0001  **Overmet need** | 54.99% |
| **Cancer** | 0.63  (0.38; 0.87)  p<0.0001 | 30% | -0.02  (-0.07; 0.04)  p=0.56  **Not significant** | 84% | -0.04  (-0.08; -0.01)  p=0.029  **Unmet need** | 23.85% |

1. McConnachie and Sutton (2004) ‘Derivation of an Adjustment to the Arbuthnott Formula for Socioeconomic Inequities in Health Care’. [↑](#footnote-ref-1)
2. The Arbuthnott index was used as a measure of deprivation in the unmet need analysis carried out by McConnachie and Sutton (2004). [↑](#footnote-ref-2)