University of Bath

Royal College of Surgeons

Healthcare Informatics

Principal Component Analysis of Practice Data from the Quality and Outcomes Framework of the United Kingdom General Practice Contract

This project is submitted in accordance with the requirements for the degree of Master of Healthcare Informatics of the Royal College of Surgeons of Edinburgh / University of Bath 2008

Dr Gavin Jamie

Supervisor: Dr Gordon Taylor

August 2008

Abstract

A new General Practice contract for United Kingdom was introduced in 2004 and introduced quality indicators which were linked to a large part of the remuneration due to practices. These indicators were known as the Quality and Outcomes Framework. The data was mostly automatically extracted from practice clinical computer systems, gathered centrally and published in England and each of the three devolved governments. It covered disease prevalence in eleven areas as well as achievement against over forty clinical targets. This was the most comprehensive data set relating to the clinical work of practices that has every been produced in the United Kingdom.

The majority of studies of this data have concentrated on the total number of points gained by each practice, either overall or in specific disease indicators. As the conversion to points is a process which results in loss of data this is less than totally satisfactory. Additionally the majority of studies have attempted to correlate this data with other health or social data sets. Using principal component analysis I attempted to find a small set of factors that would give an effective and relatively simple way of comparing practice data whilst maintaining as much of the information as possible from the original data. This was performed separately for the prevalence and achievement parts of the data.

Two factors were found for disease prevalence that explained over 70% of the variance with the first alone explaining nearly half of the total. On a simple analysis achievement data was dominated by mental health areas which dealt with small number of patients and had a consequently high variance. When areas were weighted according to their potential point score, as a surrogate measure of importance, five factors were found with high explanatory power.

These results may enhance and improve measures based on practice characteristics in future by adding a clinical aspect for the first time to these formulae with a relatively small number of simple factors. In particular resource allocation formulae for medicines and further care are currently solely based on sociodemographic data.

Contents

Chapter 1

Background

1.1 The Framework

The contract under which the majority of general practitioners (GPs) in the United Kingdom were commissioned was renegotiated and came into effect from April 2004[1]. There were many reasons that a new contract was considered necessary. Previous attempts to change the contractual arrangements with locally negotiated personal medical services (PMS) contracts had not been particularly successful with little difference in the general terms of these contracts from the standard general medical services (GMS) contracts. There was generally low morale amongst many GPs at the transfer of work from secondary care to primary care and about perceived restrictions to ways of working contained in the GMS contract. From the point of view of the government it was in the middle of a substantial increase in investment in the National Health Service and was concerned that the additional money was well spent. There was a concern amongst government ministers that there was little information about what actually happened in general practice. This was a view shared by some health economists, most publicly Alan Maynard[2] who felt that there was a black hole in relation to statistics about the quality of care in general practice. There was a declared need to change the system of remuneration from quantity of patients treated to the quality of care given - although this was obviously going to have to be translated to quantitative measures at some point to calculate payment.

There were many changes in the new contract affecting the contractual status of practices, most of them introducing greater flexibility to the delivery of the contract. The areas that received the greatest attention, perhaps unsurprisingly, were the changes to remuneration to practices. The primary source of funding for practices would be based on the practice population numbers weighted according to the Carr-Hill formula[3]. The Carr-Hill formula was designed to reflect practice workload and was based on the analysis of the consultations and populations of a sample of sixty practices. This was much criticised at the time and effectively fell out of use[4]as its effects were considered destabilising to practices. However there was also a move

towards payment based on quality measures as well as the number of patients on the practice list. This was the Quality and Outcome Framework (QOF) and could account for about 20-30% of practice income and a rather greater share of practice profits. The framework was also taken up by the majority of practices on PMS contracts.

The format eventually settled upon was one based on the collection of points for the achievement of organisational requirements and progress towards clinical targets. Points were awarded across eleven clinical areas and for organisational achievements. The points were used to calculate payments to practices. In the initial proposal for the QOF each point would have the same financial value per weighted patient but sustained pressure produced a system where the cash value of a point in a clinical area was related to practice prevalence, although not in a linear way.

Quality is, of course, somewhat hard to measure exactly. Therefore what was measured was various kinds of clinical activity that might be performed by a practice for an individual patient. This was then added up and translated into points. Thus there was not a truly qualitative assessment of a practice but rather a quantitative assessment of various indicators deemed to indicate quality in the treatment of individual patients.

This was not altogether new. In the previous version of the contract there has been various services that were paid at a set rate. Examples included contraceptive services, new patient health checks or minor surgery. What was new was a more complex way of translating the raw numbers into cash and a vastly increased range of services.

This was one of the first schemes to operate this way in the world. Much of the world looked on with interest and there are now schemes in both the USA and Australia that mirror aspects of this framework. Pay for performance has now become an established feature in the USA in both the private and Medicare sectors. However in the UK the proportion of income controlled by the scheme is higher than in other countries. Even when the amounts of cash are less there has been some concern that these incentive payments are not a cost effective use of resources. In reality the evidence base is rather thin either way[5]. Whilst the general concept has been consistent there has been some variation in how clinical judgement and patient choice has been dealt with. Whilst in the UK both upper thresholds less than 100% and exception reporting has been used other schemes have often used only one or other of these[6].

1.2 The Quality and Outcomes Framework

There were several types of targets to be achieved. Organisational areas were set up on a pass/fail basis, most of which related to having a specific service or policy in place. This perhaps the simplest to understand although it gives only binary data which is relatively uninformative about the practice. Also scores tended be very high in these areas as much of the paperwork could be shared between practices with minimal changes.

The majority of the other indicators related to clinical areas. In many of these areas there were points available for the setting up of disease registers. Again the points were awarded on a yes/no basis. However it was essential to know the size of the registers in order to calculate payment per point in each of the diseases areas.

The remainder of the areas awarded points on a sliding scale in a range of achievement against a specific indicator. The range of achievement for which payment was available started at 25% and extended to a specific figure for each indicator in the range 50-90%. Thus any achievement under 25% was not recognised and neither was any above the top of the range. None of the ranges extended over 90% - the aim being to make all points achievable. It was felt to be impossible to achieve one hundred percent due to patient factors and limitations in treatments themselves although exception reporting (see below) tackled similar issues. It has been suggested that there is some unnecessary duplication between these mechanisms and that all ranges should be extended to 100%. However in order to preserve the incentive for each patient (or the gradient of reward for each percentage point) considerable extra investment by the government would be required which is politically difficult.

The indicators varied. They included observations (e.g. measurement of blood pressure or cholesterol), an action (e.g. referral or prescription) or an outcome (e.g. target level reached of blood pressure or cholesterol). The common factor was that they were to be solidly based in evidence. However it is very possible for various organisations to look at the same evidence and come to different results. Thus the QOF indicators did not completely tally with national guidance in some areas[7, 8].

Various external factors could affect these statistics. Most commonly these were the patients themselves who may decline a specific investigation or therapy. Adverse reactions to drugs could also affect treatment or a specific treatment may also be judged to be unsuitable for specific patients for other reasons. Occasionally specific factors such as the unavailability of a test or service locally would make a certain indicator impossible to attain. In these cases specific patients could be "excepted" from specific indicators or sets of indicators. One of the most common reasons for excepting patients is an automatic mechanism where patients are excepted within a certain period of registering with a practice. Also they could be excepted from a specific area after receiving a new diagnosis. These automatic exceptions are for a period of three months for most indicators and nine months for treatment outcome indicators (e.g. blood pressure or HbA1c levels). Thus exception numbers can vary purely on practice list turnover. Exceptions have been a source of some controversy as there is considerable variation in this reporting between practices which has proved difficult to reliably explain[9].

One hundred and forty seven criteria were set across clinical and organisational areas. This included the setting up of registers of eleven diseases with the remaining clinical criteria being based on these areas. These areas covered chronic disease management only. There was nothing included about acute care and children were virtually excluded from the framework. The eleven clinical areas are given in table 1.1.

This structure was used in the years 2004/5 and 2005/6. Various changes were made for the

1. Asthma
2. Diabetes
3. Coronary Heart Disease
4. Left ventricular failure (as subset of CHD)
5. Stroke and transient ischaemic attack
6. Epilepsy
7. Mental Health
8. Chronic Obstructive Airways Disease
9. Hypothyroidism
10. Hypertension
11. Cancer

Table 1.1: Clinical areas in QOF

following years which made the data incompatible in some areas[10, 11].

1.3 The Data Generated

Data was collected separately by the Department of Health in England and health departments in each of the three devolved administrations - Scotland, Wales and Northern Ireland. In order to facilitate this automatic data extraction software was installed in the majority of practices designed around defined rules which were consistent between all four countries. This software uploaded data on the achievement in each clinical area and some organisational areas on a monthly basis to a central system. In Scotland and England this was QMAS and in Wales and Northern Ireland MSD Contract Manager was used. Only total numbers were sent. No data was sent on individual patients.

Only the data collected on the first of April was used for payment - the first event occuring in 2005. Ancillary data was calculated at other times. Prevalence was calculated on the 14th of February and the total practice population measured on the 31st of December in the previous year. This was known to practices and it is likely that practices ensured that the best possible data was presented on this date.

Once the data had been used for payment and any disputes resolved it was published four to six months after collection by each of the four countries. As the same rules were used to collect the data nationally this data was directly comparable.

As regards the clinical information two types of data were collected. For each of the clinical areas the number of patient with that diagnosis was collected. As the list size of practices

was also published this allowed a calculation of prevalence (as defined in the data collection rules) to be made. The second type of data collected relates to the achievement of targets on the populations identified. These include such areas as recording smoking history and giving advice to stop smoking. This was sent as a numerator (the number of patients with the target achieved) and a denominator (the number of patients to whom this target applies).

There is a crucial difference between these two types of data. If we assume that the data itself is perfect then it can be assumed that the prevalence data is largely determined by the prevalence of disease amongst the patients of a practice. The achievement data, however, is to be assumed to be largely under the control of the practice itself through its clinical actions.

This is unlikely to be as clear cut as these assumptions would make out as prevalence will, to a degree, be determined by the diagnostic skills and systems of the practice and achievement data by the compliance of patients. Equally the definitions of these indicators are not perfect either and may well be affected by differences in circumstance of both patients and doctors. Most obviously some of the indicators are affected by the local availability of investigative services. These two types of data also have mathematical differences. Achievement data is aiming for 100% in all areas although the practically achievable targets are more variable. Prevalence data merely tries to be accurate. There is no readily identifiable 'perfect' score.

There are limitations to the data, particularly in the latter type (achievement of targets). Patients cannot be tracked through the data. It is impossible to know whether a patient who did not receive smoking advice also did not get a cholesterol test.

QOF data has not been verified against other criteria which are considered to represent quality. Indeed as a quality measure, at least at PCT level it seem to have very little correlation with most other measures[12]. This is not perhaps entirely surprising as it was devised as a payment system and there is not gold standard quality measure in general practice to compare other criteria against.

Nevertheless this data has the strength of being nearly comprehensive across practices in the UK and Northern Ireland. Very few practices did not participate in the scheme. Those that did not participate were generally unusual PMS practices such as those catering to the homeless or patients not registered with a GP. The data is generally compiled to quite specific rules by well motivated practices. Most practices were also inspected during the first year to help confirm the data collection process.

Another significant practical advantage is that it is freely available to download. Alternative databases can cost several thousands of pounds to access. This free availability is largely the result of using data collected for another purpose, but hopefully provides compensation for that. Additionally as the data is to be made public some countries suppressed the practice level data where very small numbers of patients (typically less than six) were involved.

1.4 Types of data available

There were three types of raw data published. For much of the organisational areas there was simply a yes or no as to whether criteria had been met. Prevalence data was published for each of the eleven disease areas. This is given as a combination of the disease register size for each area and the total practice population. Finally for the clinical areas and a few organisational areas there was a numerator and denominator for calculating the ratio of patients who had met a given criteria. In most cases the denominator was based on the disease register for that disease area however it could also apply to a subset of that register. Patients may be selected by age, date of disease onset, smoking status etc. The denominators for each indicator were also affected by the application of the exception reporting.

In addition to the raw data points were also published. These were calculated directly from the points data in most cases. For most of the comparisons of QOF data between practices the total number of points achieved by each practice has used. This is relatively easy as points are directly interchangeable between indicators and disease areas. The points total is a convenient tool although there is no evidence base or validation for the ratio of points from indicator to indicator. They certainly do not represent equivalent effort or effect on disease morbidity. They are simply a stepping stone in the conversion of achievement to cash.

There is some loss of detail in the conversion to points. All prevalence data is lost as is achievement outwith the thresholds for each indicator. The high and low figures are clipped and thus some information is lost.

The final form of data was the actual cash paid to practices. Under the original contract proposals this was due to have been directly proportional to the product of points and practice list size. An additional variation was added to the clinical areas based on prevalence. The cash payment for that clinical area was adjusted by the ratio of the square root of the practice prevalence to the mean rooted prevalence calculated separately for England, Scotland, Wales and Northern Ireland. There was also an adjustment for practices whose prevalence was less than 5% of the maximum prevalence in their country. They had their prevalence increased to 5% of the maximum for the purpose of prevalence calculations.

1.5 Limitations of the data

The first rule of informatics lurks around this data[13] - this states that information should only be used for the purpose for which it was collected. The QOF result set is data that was collected as a payment mechanism. Therefore any other use has to be considered with a degree of caution. The circumstance of collection and its context have to be noted. There are clear variations noted between disease prevalence recorded by QOF and disease prevalence from other methods of collection[14]. This is to be expected but equally it is not to be assumed that one or other represents a “true” prevalence for a particular disease. In particular the definition of

mental health problems is so vague as to be practically useless.

Most of the other indicators are also difficult to find clear comparators for. The main exception is a national diabetes audit that attempted to look at similar areas both before and after the introduction of QOF[15]. Significant differences were found between the audit and QOF data for the same year which were largely ascribed to subtle differences in the rules under which the data was collected.

However much of this can be overcome by continually being aware that we are analysing the data rather than actual disease prevalence or the actual treatment given in practices. It would seem highly likely that the true disease prevalence is a significant factor in the recorded prevalence and that treatment variations are a significant factor in the reported achievement. It would be naive to assume, however, that these are the sole factors. This will also make interpretation of the produced factors easier later as we can wait until that stage before trying to ascribe meaning.

There is a second limitation to the data that is collected. Each of the indicators is measured in isolation with no cross reference between them. Only the total numbers for each area are sent from the practice. Thus it is impossible to follow patients through the system and to see whether those patients who meet the target in one area will meet the target in others also. It is also impossible to directly tell the extent of comorbidities with patients being included in more than one register. This is quite a significant limitation and not one that other databases specifically set up for research[16] also suffer from. Indeed they can produce results which show precisely the extent of comorbidities[17]. Interestingly this sort of analysis has shown that over seventy five percent of patients registered with surgeries are not included within the clinical areas of QOF.

Another potential problem is that patients may be excluded from achievement criteria if they are new to the practice or they dissent from treatment. For the first two years of the framework detailed information on exception reporting reasons was not routinely collected along with the achievement data. This was changed in Scotland and England (which used the QMAS system) but the rules under which this data was collected were not made public. If a patient was excepted for two reasons it appeared that they would only be listed once and the data was thus difficult to deal with. Additionally practice level data was released only in Scotland and much of this was suppressed due to small numbers of patients being involved in each area. In Scottish published data if an entry has fewer than six patients then that entry is suppressed in order to maintain patient confidentiality.

The nature of the data itself was also somewhat ambiguous. The data collected was comprehensive, in that all practices were looked at. Thus confidence intervals for such things as means did not seem particularly relevant. As we were dealing with all of the values in existence for a given variable confidence could be absolute. However it is also possible to look at the data as sampled patient data. Thus, for example, we can look at the subset of patients who had blood pressure measured and examine how many of them made it to target. As not all patients, even in targeted subgroups, had a measurement taken this is therefore a sampling.

There are, however, significant dangers in this approach. The sampling is not random being selected both by the actions of patients and that of the practices themselves. The figures we cannot be considered typical of the underlying patient population due to these effects which will be variable but not random. For this reason we can consider the data only descriptive of practices as entire units (patients, clinicians and systems) rather than being descriptive of the patients themselves. This is likely to vary somewhat between areas. There is less overt selection at the prevalence rather than the achievement level.

1.6 Previous work on this data

As the data has only been available since September 2005 at a national level and for a few months before that at local levels much of the work on the data is only just making it through the process to publication.

The QOF itself was running through the previous year however and there was some comment published about the effect on practice at the time[18, 19, 20, 21, 22]. This has tended to be in editorials and comment pieces and looks at the changes in consultation dynamics and practice behaviour caused by the QOF. In general this was reported in a negative way with the agenda of the patient being subverted by the agenda of the GP as directed by QOF. The effect of this is difficult to measure although it has been suggested that the 75% or so of patients who are not covered by the framework may lose out. This is unlikely to be quite as bad as the statistics suggest as those 75% are likely to be relatively underrepresented in consultation numbers and morbidity due to their being relatively more healthy than those diagnosed with chronic disease. In particular children and young people are under-represented in the data as most of the disease areas are diseases of old age.

Although the content of the targets have been based on evidence there is very little evidence that the targets themselves have influenced health outcomes in a beneficial way[23, 24]. The evidence base of QOF is for the targeted interventions, not the targets themselves. There is, in fact, relatively little evidence of changes in practices behaviour other than an increase in the coding of diseases and interventions[25].

The most common approach to the data is to compare it with other data sets and attempt to find correlations. Often the points score is used as a single summary factor although in some of the more sophisticated analyses individual clinical areas have been examined.

Other markers for morbidity have been used for comparison including hospital admission statistics[26]. In the study that measured these there was no significant co-variation between hospital admission statistics and clinical areas. This is perhaps somewhat surprising and would suggest that it cannot be assumed that results from QOF could be generalised.

There have also been attempts to correlate deprivation with various aspects of the data[24, 27, 28, 29, 26, 30, 31, 14] by multiple regression against the total QOF points score. In general these

have been unsuccessful. Even the prevalence measured in QOF of ischaemic heart disease and hypertension with alternative measures of prevalence (patient reported, population screening and secondary care observations) have not proven to be directly comparable.

Most of these studies have looked at small geographical areas. As there are not practice based statistics for most of these other medical factors then they have to be created and quite detailed information on the patient list of the practice is required making large scale research difficult. Even then it is not clear whether these constructed data sets can be considered valid as they tend to assume that populations are homogeneous within a geographical area.

The same is true for much of the social data which is used in regression analysis. This is based around national statistics, in particular census data and other measures based on this. In England social data is only available for the practice location whilst in Scotland it is available for the registered population of the practice. The use of practice location data has been shown to lead to wider variations in social data with less explanatory power of the QOF disease prevalence differences between practices[32].

Data about practitioners has been used which has revealed only modest correlation with practice size and the eduction of the physician. Again these have been looking at relatively small areas[33]. In short overall achievement is largely unpredictable based on other information.

Exception reporting has also been a source of some interest. Rates have been difficult to study as reasons for exception reporting were not collected with the other QOF data until 2006-7. Published studies have related to the years before this and have found relatively consistent exception reporting between practices but also that this has some of the most significant effects on the final calculated achievement[33]. High achieving practices tended to have the highest levels of exception reporting. Out of a maximum of 492 clinical points practices gained a median of 13.9 points or £1738 from exception reporting[6]. It was concluded that there was no evidence of a systematic use of exception reporting to defraud the system. None of these studies has attempted to look at the effects of automatic (based on registration or diagnosis date) and manual (coded) exception reporting or the effects of practice turnover although it is likely that the published data would not support a robust analysis.

More recent work has used a comparison between the first and second years to identify changes in behaviour in practices and changes over time. This may be to look for evidence of cheating the system. In a report by the Centre for Health Economics[34] there was an attempt to show patterns in the data itself, although these patterns were regressed against various other social and demographic factors around practices. In this case the main correlation was with deprivation although the aim was to find evidence to support a hypothesis of cheating by doctors. This study professed to show such evidence although this was poorly backed up by their data. This was limited to Scotland due the the difficulties in obtaining accurate social data for England discussed above.

One particular recent study stands out from some of the others by using raw data rather than points data[35]. Here the average achievement level was measured to produce a composite score.

Effectively the total scores would vary by the total variance between practices. They showed a tendency for practices in more deprived areas to have lower average levels of achievement in the first year but catching up overall by the third year. Whilst this has been presented as showing QOF reducing inequalities it may simply show practices in more deprived areas taking a little longer to comply with the targets. There is not assessment of the previous situation and the correlation with deprivation is not quite as clear cut as is suggested at times. Nevertheless there is certainly a general reduction in variation between practices over the first three years.

There have also been attempts to show differences before and after the introduction of QOF. These have been more difficult to produce as QOF was not piloted and there is no data exactly comparable. Nevertheless with a comparable and consistent method longer term analysis has been possible in diabetes at least[15, 30]. These have tended to suggest an improvement mostly in coding and the continuation of long term trends in treatment.

1.7 Factor reduction and health care data

Factor reduction techniques have been developed over the last century[36]. These techniques were initially developed for the social sciences to deal with the large amount of data produced from questionnaires or score sheets. The characteristics of these data was many different measurements which were probably not independent - one question was likely, at least in part, to be dealing with the same issues as some of the other questions. It was hoped that results could emerge from the data and define a series of hitherto unknown factors. Generally these were conceptual and difficult to measure directly. Often they were named after fairly abstract qualities such as determination, affection etc.

They slowly were used for other forms of social information including censuses. In many ways these were the same sort of data transformations although most of the concepts were more sociological than many of the previously psychological studies. Census data could be used to identify specific social groups in an objective manner. This had been done informally for years, most obviously under the class system. Objective data allowed the definition of groups which are used to describe society today, most familiarly in marketing but also in such things as deprivation scores.

Unfortunately for the uptake of this tool the calculations involved were very laborious before the advent of the digital electronic computer. A large analysis could take years to perform and could be a PhD project. This greatly limited its use to academia and large departments in government or industry. Over the final third of the 20th century automation and information technology advanced with both the production of larger and more diverse data sets and greater computational power to perform the analysis. Specific analysis of sociological data for healthcare could be undertaken and factor analysis was taken up by other fields including business and finance[37].

Medical trials had largely been performed with a single defined end point. Most were an attempt

to directly correlate a cause with disease or a treatment with outcome. Even in multifactorial conditions regression analyses against a single outcome were the norm. Large multi-centre trials could produce information on thousands of patients.

Over the past few years epidemiologists have starting looking a data sets in new ways. As we start to see new diseases then there are challenges in describing them and identifying them accurately. Perhaps the biggest new set of diagnoses is the so called metabolic syndrome. In this we see a multitude of factors including glucose tolerance, blood pressure and cholesterol measurement. It was unclear whether this was a single disease or a set of diseases with some common factors. This is not an entirely new concept - we already have sets of diseases such as cancer which have many common and overlapping factors but also distinct differences. The rise of new diagnoses in parallel with routine computerised information handling and analysis has allowed these techniques to be readily applied[38, 39, 40, 41].

Most of the published studies, however, look at data derived directly about individuals, either by questionnaire or direct measurement of individual statistics. There is very little published about data derived from organisations in the way that the Quality and Outcomes data is. In principle these organisations are no different to individuals and the absence of this sort of analysis probably reflects the tendency of information to be based on the individual with fewer publicly available multidimensional data sets about organisations.

The closest investigation to the analysis described in this dissertation is a paper describing the use of graph theory to identify patterns national health and social statistics[42]. This was explicitly geographically based and used techniques originally design for the analysis of chemical molecules. In this case a graph is considered to be a set of nodes and the connections between them rather than the charts that are commonly referred to as graphs. The geographical data sets were ideally suited to this as associations with geographical proximity and similarities of health and social data could be represented by connections and nodes respectively. It largely worked in a similar pattern to cluster analysis with larger areas of similarity being identified. This approach was described as computationally intensive, although the other methods above would probably also be considered so also. This is very much an experimental technique outside chemistry at the moment and is not currently used by any other teams for medical data.

Chapter 2

Method

2.1 Obtaining the data

The raw data was published by each of the four countries during the six months following its submission by practices at the end of March. There is a process of clarification with primary care trusts over the next few months before the final figures are passed to the information department and subsequently released and published on their websites[43, 44, 45, 46]. This happened in September and October 2006 although Scotland also published an interim and nearly complete data set in June. The data published was that which was used for payment of practices.

In all cases the data was presented on web sites as a series of Microsoft Excel spreadsheets. Each country used its own format within the spreadsheets although this has tended to be consistent from year to year. As the indicators were constant across the four countries the actual data contained was the same in each case. However a standard naming system for each indicator was no implemented and the first stage of each analysis was to apply consistent names to each of the indicators and make the format of the names identical in each case. This was achieved by converting each spreadsheet to a comma separated text file using the XLS2CSV program which is part of the CATDOC library. Next the SED stream editor was used to apply a select of replacements to each file. SED is a sophisticated programmable editor which allows complex replacements using regular expressions. For instance it could be used to apply a space in the name of an indicator between the text part and the number part and strip leading zeros in a couple of lines of code (and quite probably a single line of code if I was a better programmer).

That data was then extracted to a MySQL database by custom programs for each of the four countries. Separate programs were needed as there was no common layout to the spreadsheets between the four countries. For instance some countries prevalence data was combined with achievement data and for others this came in separate spreadsheets. The extraction system was also able to alter the format of some of the data to allow one common format in the database.

An example is the data from Scotland which gave prevalence figures as opposed to that of the other three countries which gave the register size directly. The extraction software multiplies the prevalence by the number of patients within the practice to work out the register size for the Scottish practices. Programs were written in the PHP scripting language using its command line interface.

Once the data was put into the database in a common format there can be some checks on its accuracy. In general no data was changed unless the official published data changed. There were a few exceptions to this where both the error and the correct figure were obvious. Mostly this applied to a small number of Welsh practices where the practice population was given incorrectly in a single spreadsheet. The correct figure was given in ten other places for the other ten areas. The error tended also to be obvious as the prevalence could exceed one hundred percent.

There are some missing values to the data. Some practices, although less than one percent, do not participate in the QOF. They may be highly specialised practices or have negotiated an alternative scheme. This is rare due to the effort involved in devising and managing such a scheme as opposed to the "off the shelf" solution of the national QOF.

Some of the data is deliberately withheld. This occurs because in some of the countries data is suppressed where the number of patients involved in a register or indicator is six or less. The aim in this case to preserve the anonymity of individual patients. Scotland tends to suppress the data along with Wales with England and Northern Ireland publishing all of the data. Only practices for which all data was available were used for this analysis.

The data was then exported from the database into a comma separated values format suitable to be imported into statistics programs. The data was arranged in a matrix with all of the relevant indicators in columns and the practices listed in the rows.

2.2 Analysis Method

Analysing the data from the QOF can be done in many ways. This is a relatively unknown quantity in terms of the robustness of the data and how it might be used. It is not certain how much detailed information is within the data, or indeed exactly what meaning the data carries.. The first, and possibly most obvious method of analysis is to try to show the relationship of this data to existing data sets. Much of the published work so far has show degrees of correlation with practice size, population deprivation, geography and many other factors[24, 28, 29, 31]. However all of these correlations were measured against either fractions of the QOF data, such as prevalence, or summaries such as a total of points. Additionally this has been performed in quite restricted geographical areas partly due to the difficulty in constructing comparator data sets. Typically this is across no more than a couple of PCTs whilst the largest analysis covered all of Scotland[34].

What has not been studied is any internal structure to the QOF data. Obviously dealing with 149 separate variables is difficult, but they can be reduced this to be manageable set whilst still containing more information than a simple points count. There is already a summary of the QOF data in the total points gained but this is not necessarily the best solution due to the method of calculation and the loss of detail. Also we do not have the chance to look at more than one dimension of the data when we look at the points total.

Fortunately there are several statistical tools that can be used to find structure within a data set.

In a study with a clear and defined endpoint regression analysis would be used. If were trying to use QOF data to understand their effect on, say, mortality rates this would be the way to do it. Each separate variable could be regressed against the practice mortality rate in the hope of finding those that strongly affected mortality. In this case mortality would be considered the response variable i.e. it is believed to move in response to the other variables. Regression analysis would also be possible if I was trying to look at the effect of disease prevalence on practice workload or mortality. Regression is not possible without a response variable.

In the case of the QOF however there is simply a large amount of data without any response variable to regress the others against. Other methods must be used to try to define an underlying structure to the data.

Initially cluster analysis was considered. Cluster analysis attempts to find groupings of data. To take a simple example, if we were to measure patients heights we may find two peaks in the frequency data associated with male and female patients. If we further measured voice pitch these might become more clearly clustered into two areas. Cluster analysis allows us to formalise the process of identifying these clusters. They are unlikely to be completely separate but it is enough to identify two subgroups. It was felt that although we talk of practices in cluster terms (urban, rural, suburban, inner city) it was however unlikely that these division were as clear as the descriptive language would suggest. It seems intuitively more likely that practices were distributed more in a normal pattern with a few exceptional practices. Indeed this sort of distribution would seem likely to be true along most possible axes related to engagement, practice organisation, patient population etc. Therefore continuous factors were to be looked for.

Factor reduction was the optimum method to use. This is a technique based on the fact that in most populations each variable is not completely independent. There is likely to be a degree of correlation between them to a greater or lesser extent. For instance if we measured height and weight in a population there is likely to be a degree of correlation. Equally this correlation is likely to be incomplete - some shorter people may be found to weigh as much, or indeed more than some, taller, people. We could produce a factor which we might call “size” which contained an element of both weight and height. We are likely to find that this explains more of the variation that either weight or height alone. A second, less important factor of “build” could also be included. We have therefore reduced the number of variables to one without halving the information although two factors are still needed to carry all of the information.

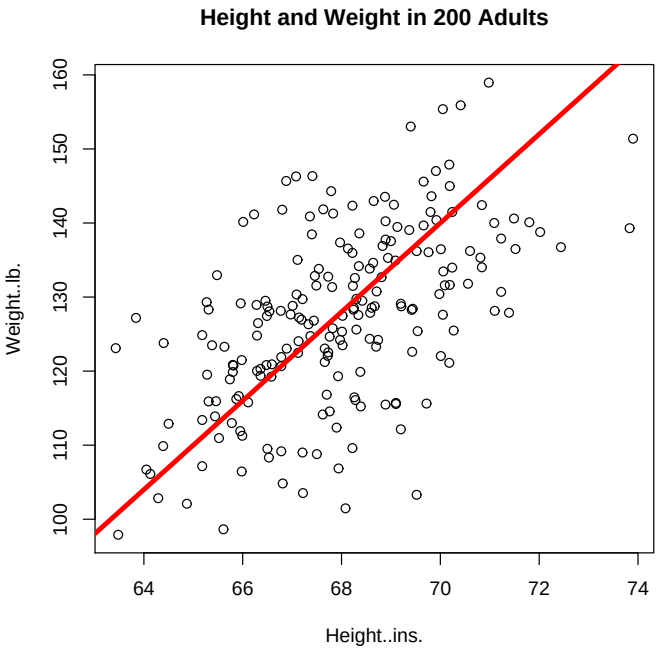

Figure 2.1: A plot of height and weight for 200 adults with a potential most significant factor illustrated in red

Some of the detail is sacrificed to reduce the complexity of the data.

A more graphical way to look at this would be to consider the general shape of the plot on a graph and an example is see in figure 2.1. As there is some correlation the plot looks roughly oval. The main factor, illustrated in red in the figure, describes the long axis of the oval and the secondary figure would describe the short axis. The main factor would describe the majority of the information. This analogy could be extended to three dimensions with a rugby ball shaped figure, or possibly a discus. The principal continues into further dimensions although their graphical manifestation is probably inconceivable.

The two main methods considered in this area were principal component analysis (PCA) and factor analysis. These are pretty similar in concept but one of the most significant differences is that PCA attempts to explain all of the variance in the data whilst factor analysis will only look to explain the covariance in the data. If there is a single area that is uncorrelated (or at least poorly correlated) with the other areas it will be ignored by factor analysis but may appear as a separate factor in PCA.

As I want the final result to be as holistic as possible it does not make a lot of sense to ignore any part of the data. If the I end up with a factor that basically mirrors one of the input variables only this is not of huge concern to me as I am looking for distinctive measures of the differences between practices. Thus if much of the variability was in a single area and the others were more tightly correlated this would be revealed by PCA.

For these reasons I chose PCA to analyse the data.

2.3 Preparing the Data

Principal component analysis will operate using either a covariance matrix or a correlation matrix as the input. There are advantages in both.

Covariance leaves the variables as they are. The covariance is strongly related to the variance of the two variables. A larger variance can dominate over a smaller one. If one area is highly variable and another less so then the covariance will tend to be similar to the larger of the two. The is particularly relevant where the variables are in different units. The covariance between weight in kg and height in metres would tend to be dominated by the weight as the mean and variance are going to be much larger. Prevalences and most of the other results from the QOF are all in the same units already, a ratio. They all already have a maximum of one. When looking at the workload of a practice sheer numbers have a very strong influence. In an average practice of 5500 patients a ten percent increase in the number of hypertensives could add another fifty patients to the register, adding ten percent to the mental health register may mean adding fewer than five patients. This does not, however fully reflect the relative workload of the two increases and the effect on the practice in general.

Precisely the opposite is true in many ways of the correlation. This effectively standardises the variables meaning that each will carry the same weight in the matrix. Standardisation is the adjusting of variables so that their means and variances are equal. Usually the mean is set at zero and the the variance at one. Correlation coefficients vary between one and minus one and will, if there is no correlation equal zero. The use of the correlation matrix makes the implicit assumption that each of the variables will affect the practice to the same degree. In prevalence this is quite unlikely to be the case but it does at least mean that the result is likely to be significantly different from just a glance at the prevalence figures. The means are really quite significantly different from one area to the next, and the variances roughly in proportion.

When achievement figures are looked at there is a much smaller range of means than in the prevalence figures. Essentially all of the figures have the same units as they are all a proportion of patients eligible to be treated. In addition it is likely that those areas with a greater variance will generally explain more about the differences between practices so a covariance matrix may be most useful. Effectively this keeps the effect of areas with very similar achievement across the country out of the final analysis (or at least relegated to one of the less significant components).

2.4 Choosing the factors

It is likely that as many factors will be generated as variables are used, whether prevalence or achievement data. Clearly factor reduction has not occurred. It is in the nature of the transformation that the first factor will explain as much of the total variance as possible and the final factor will explain very little of the variance. A decision must be made about which factors should be taken and used.

There are three main ways of doing this which are described by Dunteman[36]. Ultimately there is no correct answer of how many should be chosen. The methods mentioned here are simply suggested approaches to the process. Firstly we can aim to capture enough indicators to explain a certain amount of the total variance. A figure of 80% of the total variance has been suggested as a somewhat arbitrary threshold. This guarantees a certain amount of the variance will be explained and less than a third of the information will be discarded. If there a large number of possible factors, though, it can leave quite a lot of factors after the reduction.

A second measure is to choose all of the factors that described more than their fair share of the variance. This is most easily explained when using correlation matrices. In these cases the total variance equals the number of initial variables. Any factor with a variance of greater than one therefore explains more of the total than any one of the initial variables. The converse is also true. Again this is an arbitrary cut off and a cut off of 0.7 variance has also been suggested to retain more information. There can again be quite a lot of factors suggested by this method. In the the achievement analysis there are likely to be sixty six factors and many of them could be chosen using this method.

The final method used is a graphical one. Identifying the "elbow" of a scree plot allows the most individual factors to be selected before the start of what is generally a gentle slope as the remaining factors are broadly equal (graphically similar to a multidimensional ball). This is perhaps the least mathematical method but generally seems to produce the most useful and relevant factors. It also tends to produce fewer factors than the other methods. This latter method is the main one used although I will refer to the other methods also for comparison.

2.5 Software used

The principal component analysis was performed twice. Once on the R statistical system and once using SPSS for Windows.

The R Project for Statistical Computing[47] is an open source statistical software, originally built as an open source equivalent of the S plus statistics software. It is generally used from the command line although it is very capable of producing charts and plots in common formats. Its is available for Windows, Apple Macintosh and UNIX type systems including Linux. The versions used were those packaged as part of the Kubuntu Linux distribution. For the purpose of preparing tables for this dissertation data was exported and format in Openoffice.org Calc another open source application.

SPSS is a Windows application[48]. It is extremely extensive and expensive. A standard licence costs well over a thousand pounds. Fortunately it is available through the use of University of Bath machines. It made some sense, however, to learn an application which I had some hope of continuing to use in the future.

Chapter 3

Results

3.1 Prevalence Data

The mean and standard deviations for each of the clinical areas is given in table 3.1. It can be seen from this and the box plot (figure 3.1) that there is quite significant difference in ranges between prevalence for each area. To compare one with another is not to compare like with like. The coefficient of variation (standard deviation divided by the mean) is also shown. This gives some way of comparing the standard deviations of each area in relation to the mean. This has the same effect as standardising the means for each sample.

There is a huge range of prevalence means. The area with the greatest prevalence, hypertension, has a prevalence over thirty times that of left ventricular dysfunction. Most of the areas show a modest variation but there are some areas with greater variability. Mental health particularly falls in this category probably largely to do with the vagueness with which this area was specified in this year; it referred to as "severe and enduring mental illness" rather than specific diagnoses and could be interpreted differently for practice to practice.

The large range of standard deviations (and hence variances) confirms the need to use a correlation rather than covariance for the principal component analysis.

The first step to a Principal Component Analysis is to produce a correlation matrix for the eleven prevalence areas and this can be produced by the statistic software and is seen in table 3.2. As we are looking at a variance, not just covariance the diagonal is always one (a variable is perfectly correlated with itself!). There is a good spread of correlations from 0.71 between left ventricular dysfunction and coronary heart disease (hardly surprising as the former was a subset of the latter for the years in question) to mental health which is poorly correlated with almost everything. This again makes some sense clinically as, whilst many of the diseases listed have common risk factors such as smoking or diet, mental health does not share these. There are few very high correlations suggesting that we do not have significant redundancy in the

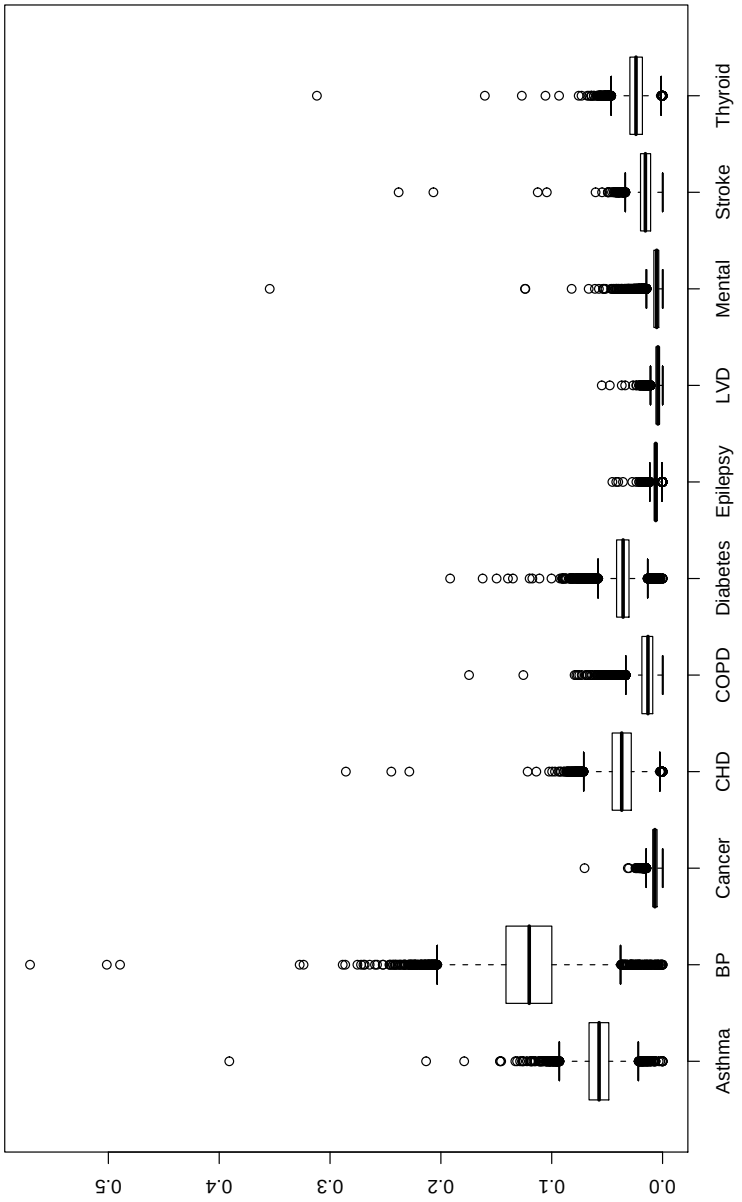

Figure 3.1: Box Plot of Prevalences

Table 3.1: Summary Data on Prevalences

Area	Mean (%)	Standard Deviation (%)	Coefficient of Variation
Asthma	5.74	1.51	0.26
Hypertension	12.06	3.50	0.29
Cancer	0.70	0.31	0.44
Coronary Heart Disease	3.70	1.39	0.37
COPD	1.47	0.88	0.60
Diabetes	3.64	1.05	0.29
Epilepsy	0.63	0.24	0.38
Left Ventricular Dysfunction	0.45	0.28	0.60
Mental Health	0.64	0.61	0.96
Stroke	1.56	0.74	0.48
Hypothyroidism	2.40	0.97	0.40

measurements i.e. two measures of prevalence do not seem to be measuring the same thing - the variables are truly independent.

We can then move on to actually calculating the factors from the matrix. As we are using correlations, discussed in the previous section, there is standardisation of the mean and the variance already. As all of the variables are independent there will be eleven factors produced.

The R function PRINCOMP(PREVALENCE,COR=1) was used. This function will generate a set of factors, the loadings of each original variable on each factor and, optionally, the factors for each of the original practices. The standard deviation of each factor is displayed as a scree plot (figure 3.2). As the total variance of the factors will equal the total variance of the original variables we can see how much of the variance is explained by each factor.

Firstly we can see from this that there is one very significant factor that accounts for almost half of the variance. The second factor is much less significant but does form quite a distinct "elbow" in the data. Only first two factors have a variance of greater than one (i.e. a greater explanatory power than the original factors). Additionally by looking at the figures the addition of the second component will take us to a total of over 55% of the variance . Looking at the loadings for the individual areas on these two factors we can see that each of the individual disease areas loads more or less on to one or both of the extracted factors.

Seeing the loadings does not actually suggest an immediate explanation for the factors that have been extracted. As a broad description cardiovascular disease (CHD, LVD, hypertension and stroke) loads most strongly onto component one although the loadings for everything bar Mental Health are between 0.25 and 0.4 (the sign of the loadings is not significant). This first component seems to relate most strongly to disease prevalence in the population in general, with an emphasis on cardiovascular disease. The second (and much less significant component) has strong and opposing correlations to Mental Health and less so to Cancer. There is not an obvious common clinical factor here - indeed what evidence there is suggests the opposite.

	Asthma	BP	Cancer	CHD	COPD	Diabetes	Epilepsy	LVD	Mental	Stroke	Thyroid
Asthma	1	0.38	0.28	0.36	0.33	0.21	0.36	0.32	0.06	0.37	0.35
BP	0.38	1	0.53	0.67	0.43	0.46	0.34	0.51	0.07	0.61	0.58
Cancer	0.28	0.53	1	0.45	0.25	0.09	0.23	0.35	0.02	0.51	0.5
CHD	0.36	0.67	0.45	1	0.66	0.42	0.5	0.71	0.09	0.75	0.6
COPD	0.33	0.43	0.25	0.66	1	0.29	0.5	0.56	0.12	0.54	0.36
Diabetes	0.21	0.46	0.09	0.42	0.29	1	0.19	0.3	0.12	0.26	0.21
Epilepsy	0.36	0.34	0.23	0.5	0.5	0.19	1	0.38	0.16	0.44	0.37
LVD	0.32	0.51	0.35	0.71	0.56	0.3	0.38	1	0.17	0.66	0.45
Mental	0.06	0.07	0.02	0.09	0.12	0.12	0.16	0.17	1	0.17	0.08
Stroke	0.37	0.61	0.51	0.75	0.54	0.26	0.44	0.66	0.17	1	0.58
Thyroid	0.35	0.58	0.5	0.6	0.36	0.21	0.37	0.45	0.08	0.58	1

Table 3.2: Correlation Matrix for Prevalence data

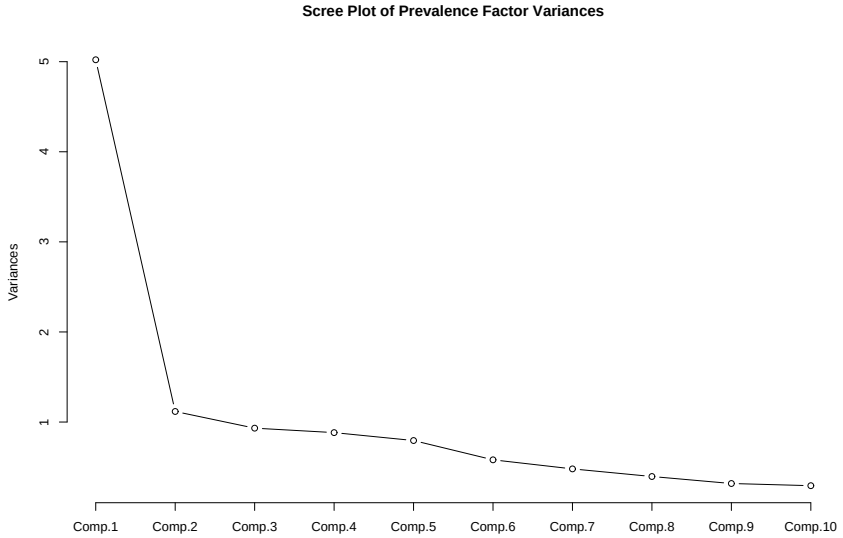

Figure 3.2: Scree Plot of Variance of Prevalence Factors

	Component One	Component Two
Asthma	-0.24	0.03
BP (hypertension)	-0.35	0.18
Cancer	-0.27	0.49
CHD	-0.4	-0.02
COPD	-0.32	-0.25
Diabetes	-0.21	-0.3
Epilepsy	-0.27	-0.26
LVD	-0.34	-0.13
Mental	-0.08	-0.64
Stroke	-0.38	0.06
Thyroid	-0.32	0.28

Table 3.3: Loadings of prevalence factors

This is likely, therefore, simply be a characteristic of populations rather than individuals.

Another way of looking at this can be seen in figure 3.3. This is a plot of all practices on the first two factors as well as arrows showing the direction of the clinical areas on these two factors. It is clearer on this diagram the separation of mental health from other prevalences as well as the clustering of the cardiovascular disease factors.

In some ways this has not been a terribly successful data reduction. We have been left with a lot of quite weak factors and the factors that are produced do not seen to be clinically useful. In choosing only two factors quite a lot of the variance is left unexplained. To get to 70% of the variance being explained would result in four factors being used. Unfortunately these are no more informative than the rest.

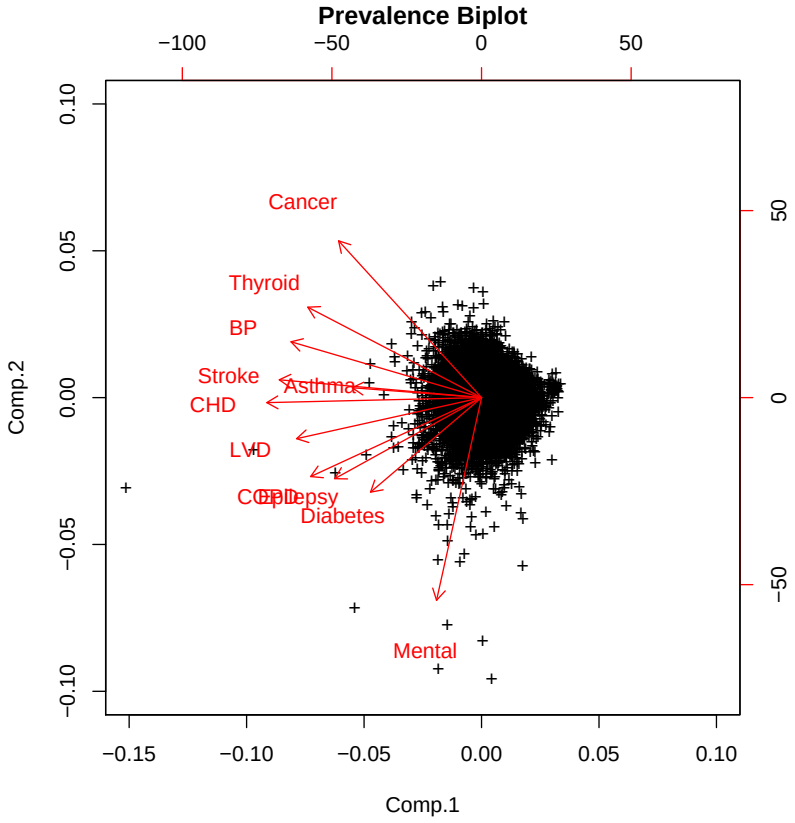

Figure 3.3: Biplot of Prevalence Factors

More positively, I have found factors that explain over 50% of the variance with just two factors. The first factor is especially significant and may well be used as a shorthand to ensure a good spread of disease prevalence in a number of practices in a simple manner.

3.2 Achievement Data

The achievement data set is substantially larger than the prevalence data with 66 different areas. Whilst this does not make a lot of difference to the mechanics of the calculation some of the presentation becomes more difficult. The correlation matrix, for instance, would cover an area thirty six times large than the one above and would be impractical to display. Table 3.4 shows the mean and standard deviation for each of the areas. All of the achievement figures are a ratio of actual versus potential achievement and thus have a minimum of zero and a maximum of one. It can seen seen fairly clearly that most of the values are within quite a small range. The maximum is 0.98 and the minimum 0.68. There is a greater range in the standard deviation with a minimum of 0.03 and a maximum of 0.28. In general those areas with the lower mean achievement have a greater standard deviation although this is probably unsurprising as there is generally more room around the mean to deviate. These are the harder areas which will tend to be the areas that demonstrate the differences between practices. The higher variance will allow them to come through in the component analysis.

Indicator	Mean	Standard Deviation
ASTHMA.7	0.73	0.12
ASTHMA.2	0.91	0.1
ASTHMA.3	0.84	0.13
ASTHMA.4	0.94	0.05
ASTHMA.5	0.89	0.09
ASTHMA.6	0.78	0.12
BP.2	0.97	0.03
BP.3	0.97	0.05
BP.4	0.92	0.05
BP.5	0.76	0.08
CANCER.2	0.88	0.19
CHD.2	0.89	0.2
CHD.3	0.96	0.05
CHD.4	0.95	0.09
CHD.5	0.97	0.05
CHD.6	0.87	0.07
CHD.7	0.93	0.07
CHD.8	0.78	0.1
CHD.9	0.94	0.06
CHD.10	0.7	0.14
CHD.11	0.88	0.13

Indicator	Mean	Standard Deviation
ASTHMA.7	0.73	0.12
CHD.12	0.91	0.08
COPD.2	0.88	0.2
COPD.3	0.89	0.17
COPD.4	0.96	0.07
COPD.5	0.95	0.11
COPD.6	0.84	0.18
COPD.7	0.91	0.14
COPD.8	0.92	0.09
CS.1	0.82	0.09
DM.2	0.94	0.07
DM.3	0.97	0.04
DM.4	0.96	0.08
DM.5	0.96	0.05
DM.6	0.62	0.11
DM.7	0.91	0.06
DM.8	0.88	0.12
DM.9	0.88	0.14
DM.10	0.87	0.15
DM.11	0.98	0.04
DM.12	0.75	0.1
DM.13	0.83	0.19
DM.14	0.95	0.06
DM.15	0.85	0.22
DM.16	0.95	0.06
DM.17	0.79	0.09
DM.18	0.9	0.08
EPILEPSY.2	0.94	0.1
EPILEPSY.3	0.94	0.11
EPILEPSY.4	0.71	0.18
LVD.2	0.89	0.25
LVD.3	0.86	0.13
MH.2	0.94	0.14
MH.3	0.9	0.26
MH.4	0.9	0.26
MH.5	0.85	0.29
STROKE.10	0.88	0.1
STROKE.2	0.86	0.22
STROKE.3	0.95	0.07
STROKE.4	0.91	0.17
STROKE.5	0.96	0.06
STROKE.6	0.85	0.1

Indicator	Mean	Standard Deviation
ASTHMA.7	0.73	0.12
STROKE.7	0.9	0.11
STROKE.8	0.72	0.13
STROKE.9	0.93	0.09
THYROID.2	0.96	0.05

Table 3.4: Variance and standard devation of achievement indicators

The areas with the greatest variance are MH 3-5. These areas are concerned with the monitoring of patients who are taking lithium. Whilst the monitoring of lithium is important - too low a dose can be ineffective, too high risks renal damage[41] - relatively few people take it. If we look at the denominator of MH 3 which should be close to the number of patients taking lithium then we find that the prevalence is a little under 1.1 per thousand patients (0.11%). Thus a typical practice of 5891 patients can expect to have six patients taking lithium. In this mythical typical practice a single patient would account for a 16% shift in achievement data - an effect that would be more pronounced still in smaller practices. It seems likely that this is one of the main reason for differences in achievement between practices in this area. The step size for one patient is very high.

The same is true of LVD 2 which relates to new diagnoses of left ventricular dysfunction. Incidence is reasonably small and variation between practices can also be large due to differences in the provision of diagnostic facilities, principally echocardiography, which are require to score against this indicator.

As we are most interested in those areas which demonstrate greatest difference between the practices standardisation was not used and a covariance matrix used for the computation of factors.

The R function PRINCOMP(ACHIEVEMENT) was used. This defaults to using a covariance matrix to calculate the factors. Again we can first look at the scree plot show to try to determine by eye the number of factors that are likely to be relevant. This is a little different to the previous scree plot as we are using covariances and so the total variance is not equal to one. In fact, and largely by coincidence, it is around 1.1 so there is only a small change in the actual numbers. We can see from the plot that around third of the total variance is accounted for in the first component, around one sixth in the second component and a twentieth in the third. Thereafter most of the remaining components are on a straight line with little to choose between them. There are sixty six components but only the first twenty are shown on the scree plot for clarity.

Encouragingly there do certainly seem to a relatively small number of significant factors. The "elbow" of the scree plot is around the third factor and looking at the cumulative variance we see that these account for the majority of the total variance. In fact only 52% of the total variance is actually included in the first three factors but to get to 70% we would have to use

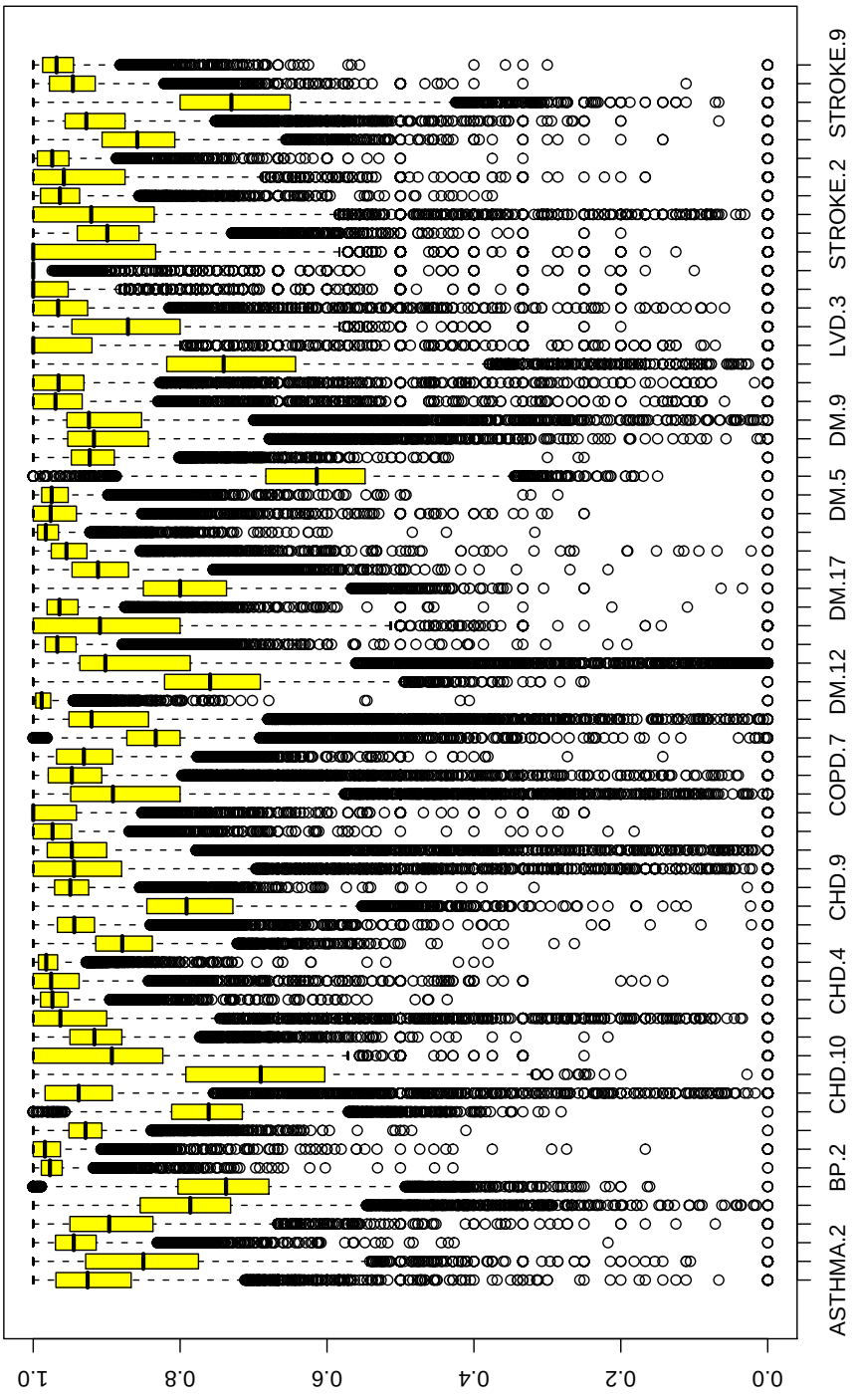

Figure 3.4: Box plot of raw achievement figures

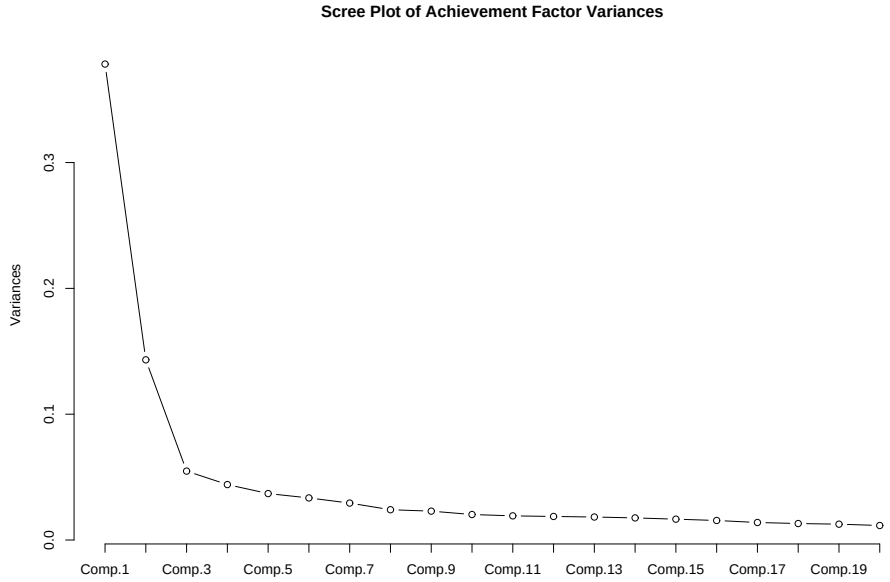

Figure 3.5: Scree Plot of Achievement Data

Area code	Description	Loading
MH 5	Lithium with recorded levels in correct range	-0.2918
LVD 2	LVD since 1/4/03 and echo confirmation	-0.2684
MH 3	Lithium who have had levels checked in the last 6 months	-0.2601
MH 4	Lithium who have creatinine and TSH checked	-0.2437
STROKE 2	Presumptive stroke since 1/4/03 and CT or MRI	-0.2350
CHD 2	Recent angina and exercise test	-0.2276
COPD 2	COPD since 1/4/2003 and have spirometry	-0.2265
COPD 6	COPD and FEV1 checked in last 27 months	-0.2012

Table 3.5: Unweighted achievement factor one

ten factors. Using all factors with greater than their share of variance would include 14 factors. Including additional factors is will not add greatly to the analysis in this case again.

We can attempt to attribute some form of meaning to these factors by looking at the loadings. The full loadings matrix for these factors is given in appendix B. Here I will list the significant loading for each of the three factors. Also included is a short description of each factor to ease interpretation.

The first factor has a negative loading for all of the indicators. This also accounts for a significant proportion of the variance. The sign of the loadings overall is arbitrary and not in any way significant. They could all be changed to positive which would changed the direction of the factor but not its significance. The significant loads are largely those with the highest original variances - which is largely to be expected given the covariance matrix used to produce the factors. Indeed the ten areas with the largest loading to this component eight of them are also in the top ten for total variance. It is, however interesting to note that all the factors are

Area code	Description	Loading
MH 5	Lithium levels in the correct range	0.5303
MH 3	Lithium who have had levels checked in the last 6 months	0.5143
MH 4	Lithium who have creatinine and TSH checked	0.4702

Table 3.6: Unweighted achievement factor two

Area code	Description	Loading
LVD 2	LVD since 1/4/03 and echo confirmation	-0.6165
STROKE 2	Presumptive stroke since 1/4/03 and CT or MRI	-0.3309
CHD 2	Recent angina and exercise test	-0.2989
DM 15	Diabetes and Proteinuria or Microalbuminuria on ACEi	-0.2580

Table 3.7: Unweighted achievement factor three

in the same direction - in this case negative. This suggests that a large part of the variance - just over a third - is explained in a "general achievement" factor.

This again is not altogether surprising. There are already figures for general achievement - the total point score. It could be hypothesised that this first factor would correlate quite strongly with the total points score. Unlike this factor, however, the total points score is weighted according the perceived difficulty of each indicator. For example the three lithium areas described above have only eleven points allocated to them - a relatively small amount. As we are looking at variances their importance is relatively exaggerated. We have, however, in this first factor eliminated quite a bit of variance and also removed the "general achievement" from the data. The remaining factors should go on to tell us a bit more about the structure of the data.

The second factor is much more strongly loaded onto the mental health lithium indicators. In fact the only positive loadings on this factor (including all of the others not listed here) are the lithium factors. This is interesting as it seems that not only is there a lot of variance in the lithium indicators but there is also a distinct axis of achievement in these indicators. There are a few areas on the opposite side of this factor, mostly related to COPD. These are interesting in themselves as these indicators were superseded by NICE guidelines[7] almost before they were introduced yet remained in the set of indicators for two further years. Anecdotally some practices tended to ignore them for this reason. If these areas could be grouped in any way they were about monitoring and reviewing - however all of their loadings were 0.11 or less. They were not about outcomes and they were not about smoking (history or intervention). They were not prescribing.

The second factor can be therefore seen as a lithium monitoring factor - perhaps against COPD and diabetes review.

The third factor has rather different loadings to the first two. The three most significant areas here relate to services that would be requested of secondary care. The descriptions in the table while quite close to the official wording are both in fact a little inaccurate. When talking about the obtaining CT scans or exercise tests they can also refer to referral to a specialist service.

These are the only three indicators of this type. They perhaps represent the availability of secondary care support for practices although they could equally well represent the efficiency of practices at recording this information in practice computer systems. The availability of specialist services may vary between different PCTs as well as local priorities and protocols. The biggest loading is against the echocardiography open access to which was is perhaps the rarest of secondary care facilities. Somewhat curious is the appearance of a prescribing indicator as the next indicator in this area. The indicator is for the prescription of angiotensin converting enzyme (ACE) inhibitors to patients with proteinuria. Even more curiously the actual testing of diabetic patients for proteinuria occurs on the other side of the factor although it can also be partially subject to the provision of an analysis service in secondary care laboratories. Other areas on the opposite site of the factor to the secondary care provision are something of a mixture but seem to be at least partially composed of asthma care in general.

We now have three factors with some explanation of the structure of the data. These have been dominated by areas with high variances. These are not necessarily important areas however. In fact the lithium areas which feature strongly in these factors affect around only a tenth of one percent of patients. Indeed, as was described previously, the small number of patients affected by these areas is likely to contribute to their relatively high variance. The factors extracted are not as useful as they might be. They are most informative about a relatively small number of patients and a small area of practice. They do not give a general picture of the difference between practices which is largely the object of the exercise in the first place. To a certain extent this problem is due to the use of a covariance matrix rather than a correlation matrix but even in a correlation matrix these areas would tend to punch above their weight. We would also dull the effect of other areas with large numbers of patients and relatively higher variance.

3.3 Weighted Achievement Data

Would it be possible, perhaps, to weight the areas according to their importance? The difficulty is in the definition of importance. We could use as a marker the number of patients affected by each one. This would give a significance based mostly on the prevalence of each disease and area although some indicators, such as smoking cessation, are based on a subset of the disease prevalence. It does not, however, allow for the difficulty of each area and we would tend to see a loss of significance for treatment and outcomes areas which are largely based on smaller denominators - for example smoking interventions. Easier areas would tend to be placed quite highly in the variance; for example measuring blood pressure is rather easier than controlling it.

We do have a subjective measure of significance in the points allocation but, as was mentioned in the introduction, there is quite a considerable degree of data loss due to the thresholds "clipping" the data. This could be modified to use the points gradient for each indicator and eliminate the effects of clipping. The points gradient is simply the rate of gain or loss of points for each indicator with the limits at either end removed. In the calculation of points for purposes of payment this gradient would only operate between the thresholds. If the thresholds

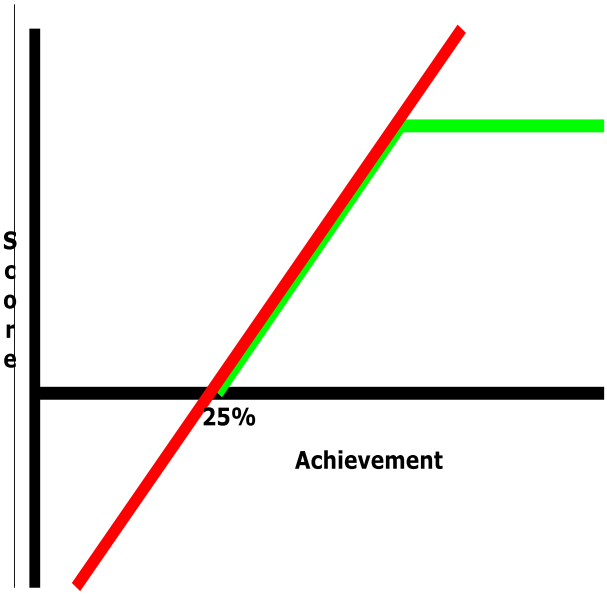

Figure 3.6: Graphical illustration of the weighting process

are removed then at the lower end the points can fall below zero and at the upper end continue until 100% achievement. Points are of course an artificial construct but the assumption is that the weighting effect of points is designed to reflect workload in each area. The truth of this assumption would be very difficult to test unfortunately - particular as the process of arriving at each point total is not well documented.

Mathematically the weighting is calculated by the formula $\frac{(A-0.25)*P}{T-0.25}$ where A is the achievement in the area as above, P is the total points available and T is the upper threshold. The 0.25 term is constant as the lower threshold for all areas in this year. Thus the information is not clipped by the process. The principle is seen in figure 3.6where the conventional points calculation is shown in green and the points gradient weighting in red. The scale simply continues for practices that have passed the top threshold. This is worth doing as in many indicators the majority of practices exceed the top threshold and so would not be differentiated with the conventional method of calculating points. It cannot be assumed that this method will always give results for any practice that are strongly related to the total point score, although in reality there is likely to be a significant degree of correlation.

It can be seen than we now have a much bigger range of standard deviations between the areas.

Area	Mean	Standard Deviation
ASTHMA.2	21.84	3.2
ASTHMA.3	7.89	1.72
ASTHMA.4	9.14	0.72
ASTHMA.5	8.47	1.26
ASTHMA.6	23.63	5.31

Area	Mean	Standard Deviation
ASTHMA.7	12.91	3.2
BP.2	11.08	0.53
BP.3	11.12	0.76
BP.4	20.68	1.54
BP.5	63.28	10.42
CANCER.2	5.84	1.72
CHD.10	12.61	3.84
CHD.11	9.77	2.05
CHD.12	7.65	0.89
CHD.2	6.87	2.21
CHD.3	7.69	0.55
CHD.4	4.33	0.58
CHD.5	7.79	0.49
CHD.6	26.21	3.02
CHD.7	7.32	0.77
CHD.8	24.19	4.67
CHD.9	7.42	0.66
COPD.2	4.83	1.5
COPD.3	4.95	1.27
COPD.4	6.57	0.64
COPD.5	6.49	1.03
COPD.6	7.85	2.38
COPD.7	6.07	1.28
COPD.8	6.68	0.87
CS.1	11.48	1.86
DM.10	2.85	0.68
DM.11	3.37	0.17
DM.12	28.48	5.64
DM.13	2.66	0.87
DM.14	3.24	0.27
DM.15	3.98	1.5
DM.16	3.24	0.25
DM.17	9.22	1.58
DM.18	3.24	0.39
DM.2	3.18	0.32
DM.3	3.34	0.19
DM.4	5.44	0.61
DM.5	3.29	0.24
DM.6	23.41	6.94
DM.7	12.11	1.18
DM.8	4.86	0.91
DM.9	2.89	0.64

Area	Mean	Standard Deviation
EPILEPSY.2	4.28	0.64
EPILEPSY.3	4.24	0.7
EPILEPSY.4	6.17	2.4
LVD.2	5.9	2.32
LVD.3	13.57	2.91
MH.2	24.24	4.78
MH.3	2.98	1.21
MH.4	3.01	1.18
MH.5	6.62	3.19
STROKE.10	2.11	0.34
STROKE.2	2.21	0.78
STROKE.3	3.24	0.34
STROKE.4	2.92	0.74
STROKE.5	2.19	0.19
STROKE.6	6.66	1.06
STROKE.7	1.99	0.33
STROKE.8	6.66	1.9
STROKE.9	4.19	0.57
THYROID.2	6.55	0.47

Table 3.8: points weighted values

We can then go through the same process with the new data set. There is considerable variation from area to area as would be expected with variations in the points total. BP 5 has the greatest mean score at 63 and stroke 10 is only 2. The very large number of points available for blood pressure control in hypertensives and the narrowness of the range (from 25-50%) has led to a very steep gradient and consequently contributes to the high standard deviation for this area. We have replaced one set of biases with another. However this is exactly the weighting that was aimed for and is rather greater than the differences between the means of the unweighted achievement indicators. These biases have had at least some consideration for relevance. The standard deviations are also tending to follow the means although by no means in a consistent fashion.

We can see that there is the shape to the scree plot (figure 3.7) that we have come to expect. There is one factor which accounts for a large proportion of the variance with another three before a distinct "elbow" and a gentle slope after this. As everything has been scaled up the variance is consequently quite a bit larger in total with the scale running into the hundreds. The first factor contains 45% of the variance with the first four factors making up just short of 70% of the total variance. Only the first eleven factors contain more than their share of the total variance.

Based on the "elbow" of the data four factors seem a reasonable number to take and should be

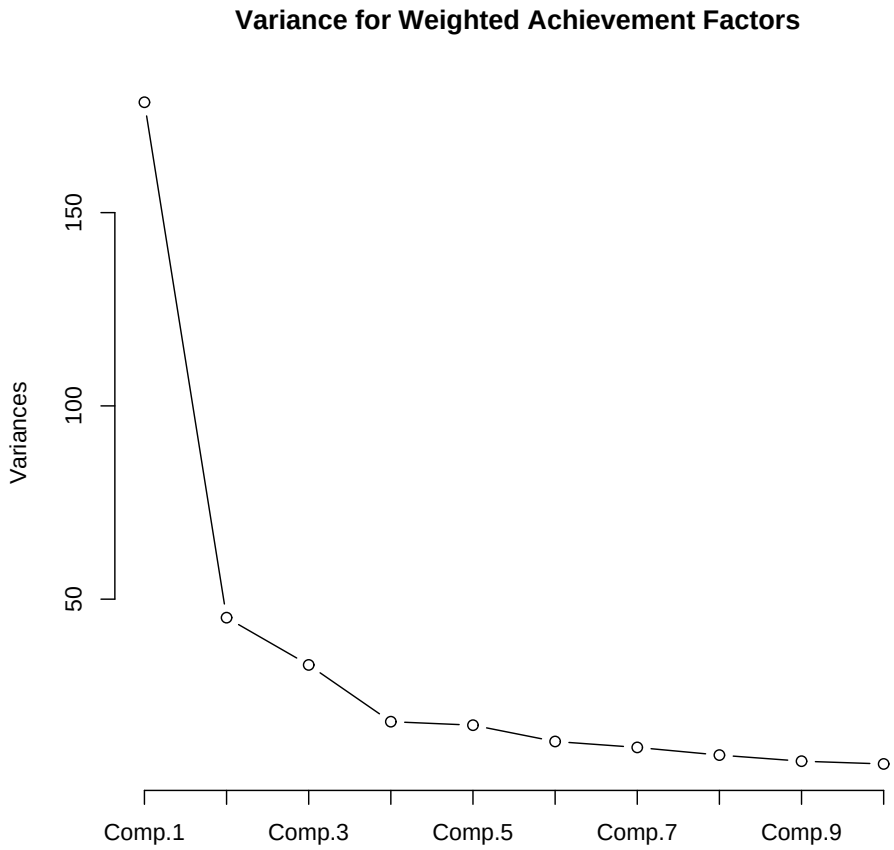

Figure 3.7: Scree Plot for weighted areas

Area code	Description	Loading
BP.5	Hypertension and BP 150/90 or less	-0.727
DM.12	Diabetes and BP 145/85 or less	-0.321
DM.6	Diabetes and HbA1c 7.4 or less	-0.255
ASTHMA.6	Asthma having review in last 15 months	-0.252
CHD.8	CHD and Cholesterol 5.0mmol/l or less	-0.224

Table 3.9: Loadings for points adjusted factor one

Area code	Description	Loading
BP.5	Hypertension and BP 150/90 or less	-0.479
DM.6	Diabetes and HbA1c 7.4 or less	0.760

Table 3.10: Loadings for points adjusted factor two

quite manageable. Again I will look at these in turn and again factors with loadings of over 0.2 are listed although these individual loadings are not likely to be important in themselves it can be a general measure of which way each of the derived factors is pulling.

The first factor is very much one of measurement. It is very strongly associated with blood pressure target achievement - not surprisingly as there are a lot of points in this area and the variance was very large. Blood pressure measurement in diabetes is also featured although less strongly. The other targets here are also dominated by outcomes measures - blood glucose levels in diabetes and cholesterol levels in coronary heart disease. Five of the top six areas are clinical outcomes measures which are fairly rare in the framework overall - despite the inclusion of the word "outcomes" in the name. These are seen as some of the more challenging indicators so have significant numbers of points allocated to them. This factor - with 45% of the variance could be described as an outcomes factor or, perhaps more accurately, treatment success. As the loading on BP 5 is so strong it is quite close to being a direct equivalent of that factor. In the strictest sense of the word outcomes would be seen as related to morbidity and mortality which is not measured in the QOF.

The second factor has a very different pattern with few significant loadings. With blood pressure outcomes on one side and HbA1c (and to a much lesser extent cholesterol) outcomes on the other it is perhaps a surprise as to the extent of the variance explained by this factor. The loading onto the HbA1c is the largest of the these and about the most significant loading here so this is again largely a factors of outcomes - largely HbA1c and blood pressure measurement. A steep points gradient tends to make HBA1c significant at the 7.4 level and a big contributor to variation in points between practices. These two areas varied together in the first factor and this second factor, with rather less explanatory power, is starting to split up the practices which varied together on the first. Once again, however, we see that much of the variance between practices is concentrated in a relatively small number of areas with large number of points.

The third factor seem to have a variety of reviews on one side with outcomes on the other. In this less significant factor HbA1c and blood pressure in diabetes are on the same side once gain, although there is a significant difference in the magnitude of their loadings. The review factors are two of only four review factors in the QOF and it is interesting that they occur

Area code	Description	Loading
ASTHMA.6	Asthma having review in last 15 months	-0.424
MH.2	Mental Health Problems reviewed in 15 months	-0.379
BP.5	Hypertension and BP 150/90 or less	0.207
DM.12	Diabetes and BP 145/85 or less	0.263
DM.6	Diabetes and HbA1c 7.4 or less	0.539

Table 3.11: Loadings for points adjusted factor three

Area code	Description	Loading
ASTHMA.6	Asthma having review in last 15 months	-0.647
ASTHMA.2	Asthma diagnosed since 1/4/2003 and PFR done	-0.223
ASTHMA.7	Asthma given influenza vaccine	-0.210
CHD.8	CHD and Cholesterol 5.0mmol/l or less	0.222
MH.5	Lithium with recorded levels in correct range	0.238
MH.2	Mental Health Problems reviewed in 15 months	0.522

Table 3.12: Loadings for points adjusted factor four

together here, especially as we would not expect a great deal of commonality between asthma and mental health areas. They are mostly an indication of getting patients "through the door" and thus would favour seeing patients more frequently or being stricter about reviews. This factor could be considered to show achievement versus process. This is certainly not a primary choice for many practices - the loadings are much less here - but rather a more subtle separation of emphasis between practices.

By the fourth factor we asthma reviews on one side and mental health on the other. It may seem a curious axis but these are areas with quite large numbers of points allocated to them so are likely to be areas with relatively strong loadings in this example. The meaning of this axis is probably not so easily expressed in clinical terms. There is not a lot of tension in clinical practice between asthma review and mental health review. However it is not compulsory for factors to have meaning within the concepts of clinical medicine. It is merely that they happen to exist.

Chapter 4

Conclusion

The aim of this process was to produce a measure, or series of factors that will allow the description of a general practice and its clinical quality data. These factors will cover both the morbidity brought to a practice by their population and the practice's response to it. It is certainly clear that these factors do exist and offer explanatory power above what would be produced from merely looking at the underlying measures. Traditional measures of completeness have been difficult to achieve in these factors and this is slightly disappointing, however these factors may still be considered as useful.

Trying to estimate the needs of a population of patients and the corresponding resources that should be allocated at practice level has been a challenge for researchers for some time. The solution in the 1980s was discussion by expert panel and the production of a formula based on social factors alone[49]. This formula was used in the remuneration of practices for around 20 years.

By 2003 more complete data sets were available but these still largely were still largely made up of social and demographic statistics. The Carr-Hill formula[3] was produced by the regression of these factors against a measure of practice workload. It was noted at the time that morbidity would be useful to measure but the statistics simply were not there. It is probably not a coincidence that the contract which introduced the Carr-Hill formula also produced the most comprehensive measure of morbidity rates (at least as perceived by practices) that has so far been seen across the UK. One major flaw to the Carr-Hill formula it was that the practice workload variable - which the others were regressed against - was not particularly reliable as a measure of need. However the uses of purely socio-demographic and no medical data in the formula is unlikely to have helped.

As can be seen from the analysis described in this paper a single, highly significant measure of morbidity would be relatively easy to incorporate into a new workload formula. Given the number of factors already contained within it and the near universal use of computers for calculation including both factors would not add significant complexity. It seems unlikely that

the social and economic factors would be replaced but the predictive value of the formula may well be enhanced by the use of clinical factors. One change would be that there would be an element of the resource formula largely under the control of the practice but the current QOF arrangements have more direct correlations between prevalence and reward this is not something that would be a large step. There is little current evidence of "gaming".

Other allocation formulae are used for other aspects of budgets. Drugs budgets have been based on prescribing units (PUs) which have been based initially on age and sex and laterally on sex and registration status (ASTRO_PU)[50]. There have been further attempts to refine the formula but these basic factors remain[51]. The addition of a disease prevalence factor may make the formula much more useful. The ASTRO-PU explains only 25% of inter practice variation in prescribing cost currently which is poor - and approaching useless in practice. The development of a new prescribing unit based partly on practice morbidity factors could improve this enormously. More work would be needed in this area to determine how much the improvement would be.

These factors may also be of use in the sampling of practices for research studies. It is already common practice to see comparisons of various aspects of practice in case and control groups in interventional studies. Mostly these are based on social and demographic features. It would be easy to see how additional factors of clinical morbidity could be added to such a comparison. Again having one or two significant factors makes this a quite feasible approach. Additionally, as QOF data is nationally and freely available it is easy for investigators to access. In fact it is significantly easier that social or demographic data which will either require specific data extraction or the payment of licensing fees.

Use of the achievement data is rather more problematic. As was discussed in section 1.6 it is difficult to find strong correlation with points data to other measures of quality in the practice. This may, of course, be down to problems with other measures of quality - this is a very difficult thing to measure in the wide context of general practice. It is however enough of a concern that trying to use these factors as a generic measurement of quality is not likely to be successful. The factors do, however, give a degree of information about how points are achieved within a practice. A very small number of factors account for a very large amount of the variance between practices. Achievement overall by practices in the QOF was considerably higher that the governments' expectations and variation between practices rather smaller than expected as many clustered near the top of the points total.

It is rather predictable that the primary factors are based around those areas with either high variance or, when using weighted factors, those with the highest points gradient. It is these areas which seem to predict overall success rather than a large generalised movement across all areas. There does not appear to be a strong single and multi-area aspect to practice quality and achievement. There is not a lot of underlying structure that is not quite apparent from simply looking at variances.

There are several limitation to the data collected. Firstly there are limitation in the data itself which were discussed in section 1.5. Principally these are that the data covers only a minority

of patients and, for the achievement data, a minority of the care that they receive. Any factors that derive are therefore going to be less than holistic. The use of other data sets may improve matters although this is unlikely to fully solve this problem. A majority of consultations may have no coded diagnosis at all[52], possibly where none was made.

Secondly we must remember that a large amount of information is thrown away when a relatively small number of factors is chosen. In the case of the achievement factors nearly half of the information has gone. For the population of practices as a whole this can be considered acceptable as a payoff for the factors that have been retained. At the level of individual practices it may be that one practice differs from others in a single factor. It is possible that one of the factors that is less significant at a national scale may be very significant in a small group of practices. Specialist practices such as those catering for universities, schools, nursing homes or the homeless may have their own factors which have been discarded.

We can thus make generalisations about practices and some of the common areas of variation but it could be dangerous to apply these factors didactically to individual practices. Nevertheless where characterisation is required this may add to the information that is available.

The final limitation that must be considered is the stability of these factors. Changes are made to the QOF on an annual basis. There are minor changes every year, and even within years. There have been more substantial changes taking effect in 2006 and 2008. Whilst many things have remained broadly similar these annual changes make year to year comparisons impossible. It is impossible to be sure that the factors that we have identified here will be stable from one year to the next. Patterns of behaviour may change. Even worse, factor loadings calculated in one year cannot be transferred to another as the underlying indicators will not be the same. The factors are thus simply a snapshot of practice - other indicators above such as ASTRO prescribing units which rely purely on demographics will not have these problems.

Even within these limitations it has been demonstrated here that meaningful factors can be generated from the data with relative ease. Indeed with these forms of analysis there does not appear to be any necessity to continue to use total point score as a simplistic summary. It seems unlikely that we should have another information resource as comprehensive about practices in the near future. It would be a significant waste not to use this data more creatively.

Bibliography

[1] New gms contract, February 2003.

[2] Alan Maynard and Karen Bloor. Trust and performance management in the medical marketplace. *J R Soc Med*, 96:532–539, November 2003.

[3] R. A Carr-Hill, N. Rice, and M. Roland. Socioeconomic determinants of rates of consultation in general practice based on fourth national morbidity survey of general practices. *BMJ*, 312:1008–12, 1996.

[4] British Medical Association. Bma - minimum practice income guarantee guidance, May 2003.

[5] Arnold M. Epstein. Pay for performance at the tipping point. *N Engl J Med*, 356:515–517, February 2007.

[6] Tim Doran, Catherine Fullwood, David Reeves, Hugh Gravelle, and Martin Roland. Exclusion of patients from pay-for-performance targets by english physicians. *N Engl J Med*, 359:274–284, July 2008.

[7] nice. Chronic obstructive pulmonary disease, June 2006.

[8] K. Short. Qof vs nice. *Br J Gen Pract*, 57:501, 2007.

[9] Sally Gainsbury. Pcts face tough questions on qof reporting. *The Health Service Journal*, pages 4–5, April 2008. PMID: 18504849.

[10] H. Lester, D. J Sharp, F. D Hobbs, and M. Lakhani. The quality and outcomes framework of the gms contract: a quiet evolution for 2006. *Br J Gen Pract*, 56:244–6, 2006.

[11] NHS Employers. Nhs employers: Contract changes for 2006/07, 2006. Following negotiations, NHS Employers and the General Practitioners Committee agreed changes to the contract for 2007/08.

[12] Pamela A Ohman-Strickland, A John Orzano, Paul A Nutting, W Perry Dickinson, Jill Scott-Cawiezell, Karissa Hahn, Michelle Gibel, and Benjamin F Crabtree. Measuring organizational attributes of primary care practices: development of a new instrument. *Health services research*, 42:1257–73, June 2007. PMID: 17489913.

[13] S. de Lusignan and C. Mimnagh. Breaking the first law of informatics: the quality and outcomes framework (qof) in the dock. *Inform Prim Care*, 14:153–6, 2006.

[14] G. ; Mackay Watt. The prevalence of coronary heart disease in general practices serving the most deprived populations in scotland and nhs greater glasgow, 2005.

[15] National diabetes audit 2004/5, 2006.

[16] I. M Carey, D. G Cook, S. De Wilde, S. A Bremner, N. Richards, S. Caine, D. P Strachan, and S. R Hilton. Developing a large electronic primary care database (doctors' independent network) for research. *Int J Med Inform*, 73:443–53, 2004.

[17] J Hippisley-Cox and M Pringle. Comorbidity of diseases in the new gms contract for gps : analysis of qresearch data, February 2005.

[18] G. Bugerem. What a load of 'qof'. *J Fam Plann Reprod Health Care*, 31:160, 2005.

[19] D. Carlisle. Quality and outcomes framework. points mean prizes for gps, but what's in it for patients? *Health Serv J*, 115:16–7, 2005.

[20] R. Greenwood, K. Shaw, and P. Winocour. Diabetes and the quality and outcomes framework: integrated care is best model for diabetes. *BMJ*, 331:1340–1340, 2005.

[21] H. Keen. Diabetes and the quality and outcomes framework: diabetes needs reintegration of primary and secondary care. *BMJ*, 331:1340, 2005.

[22] C. Kenny. Diabetes and the quality and outcomes framework. *BMJ*, 331:1097–8, 2005.

[23] R. Fleetcroft and R. Cookson. Do the incentive payments in the new nhs contract for primary care reflect likely population health gains? *J Health Serv Res Policy*, 11:27–31, 2006.

[24] L. A Sigfrid, C. Turner, D. Crook, and S. Ray. Using the uk primary care quality and outcomes framework to audit health care equity: preliminary data on diabetes management. *J Public Health (Oxf)*, 2006.

[25] Laura A. Petersen, LeChauncy D. Woodard, Tracy Urech, Christina Daw, and Supicha Sookanan. Does pay-for-performance improve the quality of health care? *Ann Intern Med*, 145:265–272, August 2006.

[26] Amy Downing, Gavin Rudge, Yaping Cheng, Yu-Kang Tu, Justin Keen, and Mark Gilthorpe. Do the uk government's new quality and outcomes framework (qof) scores adequately measure primary care performance? a cross-sectional survey of routine healthcare data. *BMC Health Services Research*, 7:166, 2007.

[27] M. Strong, R. Maheswaran, and J. Radford. Socioeconomic deprivation, coronary heart disease prevalence and quality of care: a practice-level analysis in rotherham using data from the new uk general practitioner quality and outcomes framework. *J Public Health (Oxf)*, 28:39–42, 2006.

[28] Y. Wang, C. A O'Donnell, D. F Mackay, and G. C Watt. Practice size and quality attainment under the new gms contract: a cross-sectional analysis. *Br J Gen Pract*, 56:830–5, 2006.

[29] J. Wright, D. Martin, S. Cockings, and C. Polack. Overall quality of outcomes framework scores lower in practices in deprived areas. *Br J Gen Pract*, 56:277–9, 2006.

[30] M. C Gulliford, M. Ashworth, D. Robotham, and A. Mohiddin. Achievement of metabolic targets for diabetes by english primary care practices under a new system of incentives. *Diabet Med*, 24:505–11, 2007.

[31] S. Saxena, J. Car, D. Eldred, M. Soljak, and A. Majeed. Practice size, caseload, deprivation and quality of care of patients with coronary heart disease, hypertension and stroke in primary care: national cross-sectional study. *BMC Health Serv Res*, 7:96, 2007.

[32] Gary McLean, Bruce Guthrie, Graham Watt, Mark Gabbay, and Catherine O'Donnell. Practice postcode versus patient population: a comparison of data sources in england and scotland. *International Journal of Health Geographics*, 7:37, 2008.

[33] Tim Doran, Catherine Fullwood, Hugh Gravelle, David Reeves, Evangelos Kontopantelis, Urara Hiroeh, and Martin Roland. Pay-for-performance programs in family practices in the united kingdom. *N Engl J Med*, 355:375–384, July 2006.

[34] H Gravelle, M Sutton, and A Ma. Doctor behaviour under a pay for performance contract:further evidence from the quality and outcomes framework, February 2008.

[35] Tim Doran, Catherine Fullwood, Evangelos Kontopantelis, and David Reeves. Effect of financial incentives on inequalities in the delivery of primary clinical care in england: analysis of clinical activity indicators for the quality and outcomes framework. *Lancet*, August 2008. PMID: 18701159.

[36] George H. (Henry) Dunteman. *Principal Components Analysis*. Sage Publications, Inc, July 1989.

[37] Diana Zhumabekova and Mardi Dungey. Factor analysis of a model of stock market returns using simulation-based estimation techniques, 2001.

[38] J. B Meigs. Invited commentary: insulin resistance syndrome? syndrome x? multiple metabolic syndrome? a syndrome at all? factor analysis reveals patterns in the fabric of correlated metabolic risk factors. *Am J Epidemiol*, 152:908–11; discussion 912, 2000.

[39] H. M Lakka, D. E Laaksonen, T. A Lakka, L. K Niskanen, E. Kumpusalo, J. Tuomilehto, and J. T Salonen. The metabolic syndrome and total and cardiovascular disease mortality in middle-aged men. *JAMA*, 288:2709–16, 2002.

[40] Claudio L Lafortuna, Fulvio Adorni, Fiorenza Agosti, and Alessandro Sartorio. Factor analysis of metabolic syndrome components in obese women. *Nutr Metab Cardiovasc Dis*, June 2007. PMID: 17600693.

[41] Richard Wedeen and Marc De Broe. *Nephrotoxic Metals*, page 1000. Oxford University Press, 3 edition, 2005.

[42] Peter A. Bath, Cheryl Craigs, Ravi Maheswaran, John Raymond, and Peter Willett. Use of graph theory to identify patterns of deprivation and high morbidity and mortality in public health data sets. *J Am Med Inform Assoc*, 12:630–641, November 2005.

[43] The Information Centre. The quality and outcomes framework, September 2006. Quality and outcomes framework (QOF).

[44] ISD Scotland. General practice - quality & outcomes framework, September 2006.

[45] NHS Wales. Gms contract - quality and outcomes framework, September 2006.

[46] Social Services Department of Heath and Public Safety. Quality and outcomes framework, September 2006.

[47] R Foundation for Statistical Computing. The r project for statistical computing, 2008.

[48] SPSS Inc. Spss® base 16.0 for windows®, 2008.

[49] B. Jarman. Identification of underprivileged areas. *Br Med J (Clin Res Ed)*, 286:1705–9, 1983.

[50] Darrin L Baines and David J Parry. Analysis of the ability of the new needs adjustment formula to improve the setting of weighted capitation prescribing budgets in english general practice. *BMJ : British Medical Journal*, 320, 2000. PMC1117487.

[51] D C E F Lloyd, C M Harris, and D J Roberts. Specific therapeutic group age-sex related prescribing units (star-pus): weightings for analysing general practices' prescribing in england. *BMJ*, 311:991–994, October 1995.

[52] Julia Hippisley-Cox, Mike Pringle, Ruth Cater, Alison Wynn, Vicky Hammersley, Carol Coupland, Rhydian Hapgood, Peter Horsfield, Sheila Teasdale, and Christine Johnson. The electronic patient record in primary care?regression or progression? a cross sectional study. *BMJ : British Medical Journal*, 326, June 2003. PMC162256.

Appendix A

Principal component results for disease prevalence

Component	1	2	3	4	5	6	7	8	9	10	11
Asthma	-0.4	0.03	-0.09	-0.27	0.81	0.44	-0.03	0.01	-0.01	0.04	-0.07
BP	-0.38	0.18	0.25	0.23	0.09	-0.12	0.05	0.07	-0.65	-0.52	0.08
Cancer	-0.35	0.49	-0.22	0.31	0.06	-0.09	0.64	-0.04	0.34	0.03	-0.01
CHD	-0.34	-0.02	0.1	-0.07	-0.23	0.03	-0.07	-0.01	-0.03	0.13	-0.87
COPD	-0.32	-0.25	0.01	-0.38	-0.24	0.09	0.29	0.7	0.08	-0.04	0.21
Diabetes	-0.32	-0.3	0.76	0.23	0.21	-0.17	0.09	-0.08	0.29	0.23	0.13
Epilepsy	-0.27	-0.26	-0.26	-0.43	0.13	-0.66	0.08	-0.37	-.0000986	-0.08	0.06
LVD	-0.27	-0.13	-0.01	-0.03	-0.32	0.46	-0.11	-0.47	0.32	-0.42	0.2
Mental	-0.24	-0.64	-0.43	0.59	0.13	0.02	0.03	0.12	-0.03	-0.04	-0.09
Stroke	-0.21	0.06	-0.14	0.07	-0.2	0.15	-0.07	-0.18	-0.39	0.69	0.33
Thyroid	-0.08	0.28	-0.13	0.16	0.06	-0.26	-0.69	0.31	0.33	-0.02	0.13

Table A.1: Loadings of the prevalence factors onto the disease areas

Component	Variance	Proportion of total variance	Cumulative proportion of total variance
1	5.02	46%	46%
2	1.12	10%	56%
3	0.93	8%	64%
4	0.86	8%	72%
5	0.79	7%	79%
6	0.58	5%	85%
7	0.48	4%	89%
8	0.4	4%	93%
9	0.32	3%	96%
10	0.29	3%	98%
11	0.18	2%	100%

Table A.2: Variance explained by each component

Appendix B

Unweighted achievement principal components

Component	1	2	3
ASTHMA 2	-0.07	-0.05	0.04
ASTHMA 3	-0.08	-0.08	0.13
ASTHMA 4	-0.04	-0.04	0.06
ASTHMA 5	-0.06	-0.06	0.09
ASTHMA 6	-0.11	-0.09	0.14
ASTHMA 7	-0.08	-0.08	0.14
BP 2	-0.03	-0.02	0.03
BP 3	-0.04	-0.02	0.02
BP 4	-0.05	-0.03	0.05
BP 5	-0.07	-0.06	0.1
CANCER 2	-0.19	-0.09	-0.08
CHD 10	-0.09	-0.08	0.13
CHD 11	-0.08	-0.05	-0.02
CHD 12	-0.07	-0.05	0.07
CHD 2	-0.23	-0.07	-0.3
CHD 3	-0.04	-0.03	0.04
CHD 4	-0.08	-0.05	0.03
CHD 5	-0.04	-0.02	0.03
CHD 6	-0.06	-0.05	0.07
CHD 7	-0.08	-0.04	0.06
CHD 8	-0.1	-0.06	0.09
CHD 9	-0.06	-0.04	0.03
COPD 2	-0.23	-0.11	-0.09
COPD 3	-0.19	-0.12	0.01
COPD 4	-0.06	-0.04	0.03

Component	1	2	3
COPD 5	-0.08	-0.04	0
COPD 6	-0.2	-0.13	0.05
COPD 7	-0.15	-0.1	0.06
COPD 8	-0.08	-0.05	0.05
CS 1	-0.08	-0.02	0.02
DM 10	-0.15	-0.1	0.14
DM 11	-0.03	-0.02	0.03
DM 12	-0.07	-0.06	0.11
DM 13	-0.17	-0.11	0.19
DM 14	-0.06	-0.03	0.05
DM 15	-0.16	-0.04	-0.26
DM 16	-0.06	-0.03	0.05
DM 17	-0.08	-0.05	0.1
DM 18	-0.07	-0.05	0.08
DM 2	-0.07	-0.05	0.08
DM 3	-0.04	-0.03	0.04
DM 4	-0.07	-0.04	0.05
DM 5	-0.05	-0.02	0.04
DM 6	-0.07	-0.03	0.09
DM 7	-0.06	-0.02	0.06
DM 8	-0.12	-0.06	0.07
DM 9	-0.14	-0.09	0.13
EPILEPSY 2	-0.1	-0.04	0.04
EPILEPSY 3	-0.11	-0.05	0.04
EPILEPSY 4	-0.16	-0.07	0.11
LVD 2	-0.27	-0.06	-0.62
LVD 3	-0.09	-0.05	0
MH 2	-0.11	-0.05	0.02
MH 3	-0.26	0.51	0.07
MH 4	-0.24	0.47	0.07
MH 5	-0.29	0.53	0.09
STROKE 10	-0.09	-0.06	0.06
STROKE 2	-0.23	-0.08	-0.33
STROKE 3	-0.07	-0.05	0.04
STROKE 4	-0.12	-0.04	-0.02
STROKE 5	-0.06	-0.03	0.03
STROKE 6	-0.08	-0.06	0.08
STROKE 7	-0.11	-0.07	0.06
STROKE 8	-0.13	-0.09	0.12
STROKE 9	-0.08	-0.04	0.01
THYROID 2	-0.04	-0.02	0.03

Table B.1: Loadings of the factors onto disease areas

Component	Variance	Proportion of total variance	Cumulative proportion of variance
1	0.37827	34.27%	34.27%
2	0.14319	12.97%	47.24%
3	0.05477	4.96%	52.20%
4	0.04406	3.99%	56.19%
5	0.03691	3.34%	59.53%
6	0.03345	3.03%	62.56%
7	0.02936	2.66%	65.22%
8	0.02407	2.18%	67.40%
9	0.02294	2.08%	69.48%
10	0.02026	1.84%	71.32%
11	0.01919	1.74%	73.06%
12	0.01871	1.69%	74.75%
13	0.01827	1.66%	76.41%
14	0.01757	1.59%	78.00%
15	0.01656	1.50%	79.50%
16	0.01552	1.41%	80.90%
17	0.01392	1.26%	82.16%
18	0.01309	1.19%	83.35%
19	0.01259	1.14%	84.49%
20	0.01153	1.04%	85.53%
21	0.01088	0.99%	86.52%
22	0.01069	0.97%	87.49%
23	0.01012	0.92%	88.40%
24	0.00888	0.80%	89.21%
25	0.00793	0.72%	89.93%
26	0.00766	0.69%	90.62%
27	0.00741	0.67%	91.29%
28	0.00719	0.65%	91.94%
29	0.00678	0.61%	92.56%
30	0.00665	0.60%	93.16%
31	0.00620	0.56%	93.72%
32	0.00530	0.48%	94.20%
33	0.00501	0.45%	94.66%
34	0.00464	0.42%	95.08%
35	0.00443	0.40%	95.48%
36	0.00422	0.38%	95.86%

Component	Variance	Proportion of total variance	Cumulative proportion of variance
37	0.00391	0.35%	96.21%
38	0.00371	0.34%	96.55%
39	0.00334	0.30%	96.85%
40	0.00317	0.29%	97.14%
41	0.00290	0.26%	97.40%
42	0.00255	0.23%	97.63%
43	0.00244	0.22%	97.85%
44	0.00230	0.21%	98.06%
45	0.00213	0.19%	98.25%
46	0.00175	0.16%	98.41%
47	0.00163	0.15%	98.56%
48	0.00157	0.14%	98.70%
49	0.00149	0.13%	98.84%
50	0.00141	0.13%	98.96%
51	0.00134	0.12%	99.09%
52	0.00128	0.12%	99.20%
53	0.00121	0.11%	99.31%
54	0.00110	0.10%	99.41%
55	0.00103	0.09%	99.50%
56	0.00097	0.09%	99.59%
57	0.00088	0.08%	99.67%
58	0.00070	0.06%	99.74%
59	0.00062	0.06%	99.79%
60	0.00046	0.04%	99.83%
61	0.00044	0.04%	99.87%
62	0.00042	0.04%	99.91%
63	0.00034	0.03%	99.94%
64	0.00026	0.02%	99.97%
65	0.00023	0.02%	99.99%
66	0.00014	0.01%	100.00%

Table B.2: Variance explained buy each unweighted achievement factor

Appendix C

Weighted achievement principal components

Factor	1	2	3	4
ASTHMA 2	-0.104	0.086	-0.181	-0.223
ASTHMA 3	-0.051	0.020	-0.078	-0.089
ASTHMA 4	-0.029	0.015	-0.037	-0.047
ASTHMA 5	-0.041	0.019	-0.056	-0.063
ASTHMA 6	-0.252	0.127	-0.424	-0.647
ASTHMA 7	-0.110	0.054	-0.135	-0.210
BP 2	-0.021	0.007	-0.019	0.000
BP 3	-0.021	0.014	-0.028	0.003
BP 4	-0.078	0.008	-0.055	-0.016
BP 5	-0.727	-0.479	0.207	-0.016
CANCER 2	-0.058	0.046	-0.104	0.069
CHD 10	-0.141	0.074	-0.157	0.025
CHD 11	-0.044	0.046	-0.066	0.054
CHD 12	-0.034	0.029	-0.043	0.006
CHD 2	-0.070	0.068	-0.130	0.103
CHD 3	-0.020	0.013	-0.026	0.010
CHD 4	-0.017	0.014	-0.026	0.010
CHD 5	-0.017	0.016	-0.022	0.014
CHD 6	-0.175	-0.047	-0.009	0.078
CHD 7	-0.034	0.030	-0.040	0.022
CHD 8	-0.224	0.195	-0.162	0.222
CHD 9	-0.026	0.021	-0.032	0.022
COPD 2	-0.054	0.052	-0.085	0.025
COPD 3	-0.048	0.044	-0.080	0.007

Factor	1	2	3	4
COPD 4	-0.020	0.017	-0.033	0.008
COPD 5	-0.022	0.025	-0.046	0.016
COPD 6	-0.091	0.076	-0.151	-0.004
COPD 7	-0.047	0.037	-0.083	-0.018
COPD 8	-0.028	0.025	-0.042	0.001
CS 1	-0.052	0.061	-0.077	0.028
DM 10	-0.028	0.018	-0.026	0.001
DM 11	-0.006	0.007	-0.004	0.002
DM 12	-0.321	-0.125	0.263	0.129
DM 13	-0.031	0.027	-0.025	0.005
DM 14	-0.011	0.014	-0.007	0.004
DM 15	-0.028	0.042	-0.033	0.041
DM 16	-0.011	0.013	-0.007	0.003
DM 17	-0.072	0.080	0.004	0.036
DM 18	-0.015	0.014	-0.013	-0.004
DM 2	-0.014	0.011	-0.011	-0.002
DM 3	-0.007	0.005	-0.006	0.001
DM 4	-0.019	0.015	-0.022	0.008
DM 5	-0.010	0.012	-0.006	0.004
DM 6	-0.255	0.760	0.539	-0.093
DM 7	-0.050	0.093	0.011	0.012
DM 8	-0.035	0.030	-0.033	0.010
DM 9	-0.026	0.017	-0.025	0.001
EPILEPSY 2	-0.022	0.015	-0.036	0.018
EPILEPSY 3	-0.025	0.017	-0.041	0.019
EPILEPSY 4	-0.083	0.059	-0.091	0.052
LVD 2	-0.064	0.073	-0.118	0.107
LVD 3	-0.072	0.067	-0.110	0.076
MH 2	-0.153	0.153	-0.379	0.522
MH 3	-0.020	0.037	-0.044	0.085
MH 4	-0.019	0.034	-0.040	0.072
MH 5	-0.061	0.101	-0.114	0.238
STROKE 10	-0.012	0.009	-0.014	0.003
STROKE 2	-0.025	0.022	-0.042	0.032
STROKE 3	-0.012	0.007	-0.014	0.007
STROKE 4	-0.016	0.015	-0.026	0.012
STROKE 5	-0.007	0.005	-0.007	0.004
STROKE 6	-0.053	-0.013	-0.006	0.017
STROKE 7	-0.014	0.009	-0.015	0.007
STROKE 8	-0.085	0.058	-0.053	0.052
STROKE 9	-0.017	0.014	-0.021	0.012
THYROID 2	-0.015	0.012	-0.013	0.001

Factor	1	2	3	4

Table C.1: Loadings of weighted factors to disease areas

Component	Variance	Proportion of total variance	Cumulative Proportion
1	178.56	45.16%	45.16%
2	45.21	11.43%	56.60%
3	32.99	8.34%	64.94%
4	18.3	4.63%	69.57%
5	17.4	4.40%	73.97%
6	13.18	3.33%	77.30%
7	11.65	2.95%	80.25%
8	9.66	2.44%	82.69%
9	8.08	2.04%	84.74%
10	7.37	1.86%	86.60%
11	6.68	1.69%	88.29%
12	5.57	1.41%	89.70%
13	4.09	1.03%	90.73%
14	3.77	0.95%	91.68%
15	3.73	0.94%	92.63%
16	3.11	0.79%	93.41%
17	2.62	0.66%	94.08%
18	2.45	0.62%	94.70%
19	2.2	0.56%	95.25%
20	1.86	0.47%	95.72%
21	1.78	0.45%	96.17%
22	1.61	0.41%	96.58%
23	1.44	0.36%	96.94%
24	1.19	0.30%	97.25%
25	0.97	0.24%	97.49%
26	0.88	0.22%	97.71%
27	0.81	0.21%	97.92%
28	0.67	0.17%	98.09%
29	0.65	0.17%	98.25%
30	0.61	0.15%	98.41%
31	0.52	0.13%	98.54%
32	0.5	0.13%	98.67%
33	0.44	0.11%	98.78%
34	0.4	0.10%	98.88%
35	0.38	0.10%	98.98%
36	0.38	0.10%	99.07%
37	0.34	0.09%	99.16%
38	0.32	0.08%	99.24%

Component	Variance	Proportion of total variance	Cumulative Proportion
39	0.3	0.08%	99.32%
40	0.28	0.07%	99.39%
41	0.27	0.07%	99.46%
42	0.25	0.06%	99.52%
43	0.22	0.06%	99.57%
44	0.21	0.05%	99.63%
45	0.18	0.05%	99.67%
46	0.17	0.04%	99.72%
47	0.15	0.04%	99.75%
48	0.14	0.03%	99.79%
49	0.13	0.03%	99.82%
50	0.12	0.03%	99.85%
51	0.12	0.03%	99.88%
52	0.08	0.02%	99.90%
53	0.08	0.02%	99.92%
54	0.06	0.01%	99.93%
55	0.05	0.01%	99.95%
56	0.04	0.01%	99.96%
57	0.04	0.01%	99.96%
58	0.03	0.01%	99.97%
59	0.03	0.01%	99.98%
60	0.02	0.01%	99.99%
61	0.02	0.01%	99.99%
62	0.01	0.00%	99.99%
63	0.01	0.00%	100.00%
64	0.01	0.00%	100.00%
65	0.01	0.00%	100.00%
66	0	0.00%	100.00%

Table C.2: Variance explained by each of the weighted factors

Appendix D

The clinical indicator definitions

Code	Definition
ASTHMA 1	The practice can produce a register of patients with asthma, excluding patients with asthma who have been prescribed no asthma-related drugs in the last twelve months
ASTHMA 2	The percentage of patients aged eight and over diagnosed as having asthma from 1 April 2003 where the diagnosis has been confirmed by spirometry or peak flow measurement
ASTHMA 3	The percentage of patients with asthma between the ages of 14 and 19 in whom there is a record of smoking status in the previous 15 months
ASTHMA 4	The percentage of patients aged 20 and over with asthma whose notes record smoking status in the past 15 months, except those who have never smoked where smoking status should be recorded only once since diagnosis
ASTHMA 5	The percentage of patients with asthma who smoke, and whose notes contain a record that smoking cessation advice or referral to a specialist service, where available, has been offered within the last 15 months
ASTHMA 6	The percentage of patients with asthma who have had an asthma review in the last 15 months
ASTHMA 7	The percentage of patients aged 16 and over with asthma who have had influenza immunisation in the preceding 1 September to 31 March
BP 1	The practice can produce a register of patients with established hypertension
BP 2	The percentage of patients with hypertension whose notes record smoking status at least once

Code	Definition
BP 3	The percentage of patients with hypertension who smoke, whose notes contain a record that smoking cessation advice or referral to a specialist service, if available, has been offered at least once
BP 4	The percentage of patients with hypertension in whom there is a record of the blood pressure in the past 9 months
BP 5	The percentage of patients with hypertension in whom the last blood pressure (measured in the last 9 months) is 150/90 or less
CANCER 1	The practice can produce a register of all cancer patients diagnosed after 1 April 2003
CANCER 2	The percentage of patients with cancer diagnosed from 1 April 2003 with a review by the practice recorded within six months of confirmed diagnosis. This should include an assessment of support needs, if any, and a review of co-ordination arrangements with secondary care
CHD 1	The practice can produce a register of patients with coronary heart disease
CHD 2	The percentage of patients with newly diagnosed angina (diagnosed after 1 April 2003) who are referred for exercise testing and/or specialist assessment
CHD 3	The percentage of patients with coronary heart disease whose notes record smoking status in the past 15 months, except those who have never smoked where smoking status need be recorded only once since diagnosis
CHD 4	The percentage of patients with coronary heart disease who smoke, whose notes contain a record that smoking cessation advice or referral to a specialist service, where available, has been offered within the last 15 months
CHD 5	The percentage of patients with coronary heart disease whose notes have a record of blood pressure in the previous 15 months
CHD 6	The percentage of patients with coronary heart disease in whom the last blood pressure reading (measured in the last 15 months) is 150/90 or less
CHD 7	The percentage of patients with coronary heart disease whose notes have a record total cholesterol in the previous 15 months
CHD 8	The percentage of patients with coronary heart disease whose last measured total cholesterol (measured in last 15 months) is 5 mmol/l or less
CHD 9	The percentage of patients with coronary heart disease with a record in the last 15 months that aspirin, an alternative anti-platelet therapy, or an anti-coagulant is being taken (unless a contraindication or side-effects are recorded)

Code	Definition
CHD 10	The percentage of patients with coronary heart disease who are currently treated with a beta blocker (unless a contraindication or side-effects are recorded)
CHD 11	The percentage of patients with a history of myocardial infarction (diagnosed after 1 April 2003) who are currently treated with an ACE inhibitor or angiotensin II antagonist
CHD 12	The percentage of patients with coronary heart disease who have a record of influenza immunisation in the preceding 1 September to 31 March
COPD 1	The practice can produce a register of patients with COPD
COPD 2	The percentage of patients in whom diagnosis has been confirmed by spirometry including reversibility testing for newly diagnosed patients with effect from 1 April 2003
COPD 3	The percentage of all patients with COPD in whom diagnosis has been confirmed by spirometry including reversibility testing
COPD 4	The percentage of patients with COPD in whom there is a record of smoking status in the previous 15 months, except those who have never smoked where smoking status need be recorded only once since diagnosis
COPD 5	The percentage of patients with COPD who smoke, whose notes contain a record that smoking cessation advice or referral to a specialist service, where available, has been offered in the past 15 months
COPD 6	The percentage of patients with COPD with a record of FeV 1 in the previous 27 months
COPD 7	The percentage of patients with COPD receiving inhaled treatment in whom there is a record that inhaler technique has been checked in the preceding 27 months
COPD 8	The percentage of patients with COPD who have had influenza immunisation in the preceding 1 September to 31 March
DM 1	The practice can produce a register of all patients with diabetes mellitus
DM 2	The percentage of patients with diabetes whose notes record BMI in the previous 15 months
DM 3	The percentage of patients with diabetes in whom there is a record of smoking status in the previous 15 months, except those who have never smoked where smoking status need be recorded only once since diagnosis
DM 4	The percentage of patients with diabetes who smoke and whose notes contain a record that smoking cessation advice or referral to a specialist service, where available, has been offered in the last 15 months

Code	Definition
DM 5	The percentage of diabetic patients who have a record of HbA 1c or equivalent in the previous 15 months
DM 6	The percentage of patients with diabetes in whom the last HbA 1C is 7.4 or less (or equivalent test/reference range depending on local laboratory) in last 15 months
DM 7	The percentage of patients with diabetes in whom the last HbA 1C is 10 or less (or equivalent test/reference range depending on local laboratory) in last 15 months
DM 8	The percentage of patients with diabetes who have a record of retinal screening in the previous 15 months
DM 9	The percentage of patients with diabetes with a record of the presence or absence of peripheral pulses in the previous 15 months
DM 10	The percentage of patients with diabetes with a record of neuropathy testing in the previous 15 months
DM 11	The percentage of patients with diabetes who have a record of the blood pressure in the past 15 months
DM 12	The percentage of patients with diabetes in whom the last blood pressure is 145/85 or less
DM 13	The percentage of patients with diabetes who have a record of micro-albuminuria testing in the previous 15 months (exception reporting for patients with proteinuria)
DM 14	The percentage of patients with diabetes who have a record of serum creatinine testing in the previous 15 months
DM 15	The percentage of patients with diabetes with a diagnosis of proteinuria or micro-albuminuria who are treated with ACE inhibitors (or A2 antagonists)
DM 16	The percentage of patients with diabetes who have a record of total cholesterol in the previous 15 months
DM 17	The percentage of patients with diabetes whose last measured total cholesterol within the previous 15 months is 5mmol/l or less
DM 18	The percentage of patients with diabetes who have had influenza immunisation in the preceding 1 September to 31 March
EPILEPSY 1	The practice can produce a register of patients receiving drug treatment for epilepsy
EPILEPSY 2	The percentage of patients aged 16 and over on drug treatment for epilepsy who have a record of seizure frequency in the previous 15 months
EPILEPSY 3	The percentage of patients aged 16 and over on drug treatment for epilepsy who have a record of medication review in the previous 15 months

Code	Definition
EPILEPSY 4	The percentage of patients aged 16 and over on drug treatment for epilepsy who have been seizure free for the last 12 months recorded in the last 15 months
LVD 1	The practice can produce a register of patients with CHD and left ventricular dysfunction
LVD 2	The percentage of patients with a diagnosis of CHD and left ventricular dysfunction (diagnosed after 1 April 2003) which has been confirmed by an echocardiogram
LVD 3	The percentage of patients with a diagnosis of CHD and left ventricular dysfunction who are currently treated with ACE inhibitors (or A2 antagonists)
MH 1	The practice can produce a register of people with severe long-term mental health problems who require and have agreed to regular follow-up
MH 2	The percentage of patients with severe long-term mental health problems with a review recorded in the preceding 15 months. This review includes a check on the accuracy of prescribed medication, a review of physical health and a review of co-ordination arrangements with secondary care
MH 3	The percentage of patients on lithium therapy with a record of lithium levels checked within the previous 6 months
MH 4	The percentage of patients on lithium therapy with a record of serum creatinine and TSH in the preceding 15 months
MH 5	The percentage of patients on lithium therapy with a record of lithium levels in the therapeutic range within the previous 6 months
STROKE 1	The practice can produce a register of patients with stroke or TIA
STROKE 2	The percentage of new patients with presumptive stroke (presenting after 1 April 2003) who have been referred for confirmation of the diagnosis by CT or MRI scan
STROKE 3	The percentage of patients with TIA or stroke who have a record of smoking status in the last 15 months, except those who have never smoked where smoking status need be recorded only once since diagnosis
STROKE 4	The percentage of patients with a history of TIA or stroke who smoke and whose notes contain a record that smoking cessation advice or referral to a specialist service, if available, has been offered in the last 15 months
STROKE 5	The percentage of patients with TIA or stroke who have a record of blood pressure in the notes in the preceding 15 months

Code	Definition
STROKE 6	The percentage of patients with a history of TIA or stroke in whom the last blood pressure reading (measured in last 15 months) is 150/90 or less
STROKE 7	The percentage of patients with TIA or stroke who have a record of total cholesterol in the last 15 months
STROKE 8	The percentage of patients with TIA or stroke whose last measured total cholesterol (measured in last 15 months) is 5 mmol/l or less
STROKE 9	The percentage of patients with a stroke shown to be non-haemorrhagic, or a history of TIA, who have a record that aspirin, an alternative anti-platelet therapy, or an anti-coagulant is being taken (unless a contraindication or side-effects are recorded)
STROKE 10	The percentage of patients with TIA or stroke who have had influenza immunisation in the preceding 1 September to 31 March
THYROID 1	The practice can produce a register of patients with hypothyroidism
THYROID 2	The percentage of patients with hypothyroidism with thyroid function tests recorded in the previous 15 months