HSE
Health & Safety
Executive

Development of statistical approaches to the handling and analysing of large occupational data sets

Prepared by **Dr. John M. Thompson**
for the Health and Safety Executive 2005

RESEARCH REPORT 373

Development of statistical approaches to the handling and analysing of large occupational data sets

Dr. John M. Thompson
Consultant Statistician
B.Sc., M.Sc., Ph.D.,
49 Hazel Drive
Armitage
Rugeley
Staffordshire
WS15 4TZ

Methods were developed to extract and prepare data from large complex occupational exposure databases, to undertake statistical explorations and analyses relating to possible changes in exposure patterns that may have been influenced by the impact of legislation and changes in regulatory exposure limits. A database compiled from exposure measurements collected over a 30-year period, primarily from iron and steel foundries throughout the UK, was used as the example with which to develop and test the statistical methods. These methods were then applied to develop and test hypotheses about the impact on exposure patterns of respirable silica and dust. These explorations and analyses showed that there were significant reductions in exposures to these hazardous substances, especially in the upper exposure ranges, demonstrating the valuable impact of both legislation and the tightening of regulatory limits and the importance of monitoring in demonstrating this sustained impact in reducing exposure risks.

This report and the work it describes were funded by the Health and Safety Executive. Its contents, including any opinions and/or conclusions expressed, are those of the author alone and do not necessarily reflect HSE policy.

This report makes reference to work carried out by Castings Technology International that was also funded by HSE. That work is published as Research Report 374 (2005) Evaluating the Effectiveness of Legislation, Technology and Working Methods for Reducing Occupational Exposure in the Foundry Industry.

HSE BOOKS

First published 2005

ISBN 0 7176 6150 4

Acknowledgements

I am most grateful to the following Health and Safety Executive staff for valuable discussions during the period in which this project was being planned and also during the course of the project itself: Dr. Andy Phillips, Dr. Celia Elliott-Minty, Mr. Damien McElvenny.

Dr. Glynn Morley, Mr. David Wells and Mr. Andrew Greenall from CTi were involved in compiling the occupational exposure database covering a thirty-year period in a large range of UK foundries. They were all very helpful in explaining the various categories of jobs and processes within the foundry industry and the methods of sampling and measurement used in collecting the data entered into the database. I am especially grateful to Mr. Andrew Greenall for extracting linked exposure data to enable me to explore relationships between hazardous substances sampled simultaneously in many of their exposure monitoring exercises.

CONTENTS

EXECUTIVE SUMMARY

The objectives of work reported here were to develop methodologies (i) for interrogating large occupational exposure databases and extracting relevant sets of data and associated metadata (data about that data) and (ii) to develop appropriate statistical methodologies for exploration and analysis of extracted data and metadata. Castings Technology International (CTi) was contracted by HSE to compile a large database of occupational exposures from the UK foundry industry covering a thirty year period. That database was used in my work as the example database and was compiled in MS Excel. Data was extracted from it using the Advanced Filter function and, for linked data, using Pivot Tables. These methods were used to create MS Excel files of subsets of the database that could be imported via ODBC (Open Data Base Connectivity) into the Statistical Software package Minitab v13 for subsequent data preparation, exploration and analysis. I also provided core support to CTi and collaborated with them in applying statistical methodologies developed to (i) validate data quality (ii) undertake data analysis and interpretation and (iii) provide advice, guidance and training to assist them to interpret data in the database, including validation and checking of entries during database compilation. I suggested modifying database structure to better enable subsequent extraction of the data and associated metadata, linked by being collected at the same time by the same device. During this editorial activity, concerns were raised about reliability of some respirable silica assays done at a particular laboratory. I did exploratory statistical analyses to evaluate sample data structure collected during this period in comparison with contemporary data obtained from assays performed by another laboratory whose assay methods were considered more reliable. I advised CTi, as a result, to remove the suspect data from their database.

I examined various reliable subsets of the database, especially for respirable dust and silica (because of HSE’s interest in these). I used robust/resistant exploratory and nonparametric statistical methods, several of which I adapted to enable testing of hypotheses generated in the initial project development, while others were generated as a result of my initial exploratory analyses. Standard parametric statistical methods were not used because no sample data in the database conformed to standard theoretical probability distribution. Various hypotheses were formulated around the expectation that, if Health and Safety legislation and reductions in regulatory exposure limits were valuable instruments in risk reduction, then significant downward trends with time should be detectable, linked with the time-line of the introduction of law and regulation, with lags related to responses to the new requirements. Hypotheses were focussed around the idea that the greatest impact of regulation would be on reduction in higher exposure levels, so changes in sample data distribution shape would be expected. The most important tool in starting such shape change evaluation in annual exposure patterns was Letter Value Analysis, which allows detailed examination of the distribution pattern when progressing outwards from the sample median into the tails. Upper Letter Values (Fourths, Eighths and Sixteenths corresponding to specific upper percentiles: 75, 87.5 and 93.75) were estimated and used in subsequent analyses either as testable parameters or as lower thresholds for further data sub-setting and subsequent analysis. The most useful methods adapted for this further analysis were Kruskal Wallis One Way Analysis of Variance (ANOVA) by Ranks, parallel box and whisker plots with confidence intervals on the medians (a graphical form of ANOVA), Lowess Smoothing Plots and Cusum Span Plots. These allowed testing of when changes in sample data distribution shape occurred and to what extent. Graphical tools so used are of great value in conveying, to non-specialists and specialists alike, the magnitude of any changes that might have occurred.

Very significant reductions (70 to 94%) were found to have occurred in all upper levels of exposures to respirable dusts and silica over the two decades since introduction of the Health and Safety at Work, etc. Act. Concern is raised about recent drops in monitoring activity that could be used to verify sustainability of these desirable reductions in occupational exposures.

CHAPTER ONE
INTRODUCTION AND OBJECTIVES

Introduction

In this research project, I have developed an approach to the use of statistical methodology that is designed to enable trends in large industrial hygiene databases to be revealed and so compared to appropriate timelines of changes in Health and Safety legislation as well as in regulatory limits. In this way, it becomes possible to explore the relationships between these legislative and regulatory changes and the changes that may have occurred in occupational exposure patterns.

What I have found fascinating and rewarding about this work is that it has revealed some genuine beneficial impacts on occupational exposures that seem to be strongly linked along the historical time-line with the introduction of Health and Safety legislation and to changes in regulatory occupational exposure limits. I have illustrated the data analysis methodology with some important examples of such beneficial impacts. I hope that this will stimulate further investigations applying this methodology.

Being heavily involved in Healthcare Technology Assessment (HTA), I would like to draw the reader’s attention to currently accepted definitions of HTA (see NHS HTA Reports on NHS website and Kristensen FB et al, 2001). Occupational Exposure Monitoring and its assessment are certainly Healthcare Technologies and as such their impact and value should be subject to rigorous assessment using the suite of advanced HTA methodologies currently being applied to other healthcare technologies.

Industry should undertake occupational exposure monitoring as part of its practice of good governance and as an obligation under Health and Safety law. It is my view that the Health and Safety Executive also should have such an obligation under its draft Public Service Agreement (PSA) linked to the UK Government’s Revitalising Health and Safety (RHS) Targets. One cannot ignore the potential benefit of the public health impact of this PSA, along with the benefits to the UK economy as a whole. Such an HTA investigation would require not merely the conventional health economic analysis but a more innovative approach bringing together biomedical engineering, applied statistics (especially in the area of diagnostic and prognostic medical tests), public health epidemiology and the types of health economics used in HTA. It could initially be applied to a specific occupational exposure issue of major importance as a feasibility study in order to explore the most appropriate approaches to this new kind of Health Technology Assessment.

The methodological approach that I have developed enables the significance of trends or other forms of change in exposure patterns to be evaluated. It also enables the comparison of exposure patterns over various periods or blocks of time amongst different occupational tasks, industrial processes and sizes of industrial organisation undertaking those tasks and processes.

This approach has been developed using subsets of data from the database that has been prepared by Castings Technology International in a parallel HSE-funded research project. That database consists of occupational exposure data collected over a period of about 30 years primarily from the UK Iron and Steel Foundry Industry. Their database is intended to be incorporated at some stage into the HSE National Exposure Database.

In parallel with the development of the statistical methodological approach, I have collaborated with staff at Castings Technology International (CTi) in the development of methods of extracting subsets of data from their database. I have subsequently used various kinds of data preparation methodology primarily within the statistical software Minitab so as to prepare that data for statistical investigation. This data preparation process is crucial to achieving successful and reliable data analysis as it necessarily incorporates data quality management as well as elements of data validation.

Objectives

The principal objective has been to develop methodologies for statistical investigation of such data that are within reach of HSE staff and others with interests in exploring exposure patterns in a sound manner and that are not too difficult to learn and then to use these in new investigations.

Although the statistical software used in this project has been Minitab v13, the reader is also directed to other statistical software that may equally well be used for this purpose. It should be pointed out that some of the techniques applied in this work can be used by hand without the need for software but merely the assistance of a hand-held electronic calculator. This is particularly so with small to modest sized data sets.

I am also concerned that the reader should appreciate both the strengths and weaknesses of these methodologies. It should be worthwhile to learn to apply them in a constructively critical manner, not merely as a “black box” of tools. Then the user is being empowered to extract genuinely useful information from occupational exposure data and the metadata (data about data) associated with it.

I hope that readers of this report will be motivated to practice the methodology offered, using some of the data sets that I have worked upon, thereby developing confidence in applying these tools in other areas of interest to themselves.

CHAPTER TWO
CHARACTERISTICS OF OCCUPATIONAL EXPOSURE DATA AND THEIR IMPACT ON APPROACHES TO DATA EXPLORATION AND ANALYSIS

The shape of the pattern of the data that may be collected is the shape of what is otherwise known as that of the sample population distribution. Some might ask: “Is this sample population representative of the underlying population distribution?” That is the conventional kind of statistical question about samples of observational data collected “in the field”. Another important statistical question is relevant in Industrial Hygiene and Occupational Health:
“Is this sample of data appropriate in enabling us to meet the objectives of our monitoring exercise and enable us to ask the questions that it was designed to try and answer (or to test hypotheses that the exercise was designed to test)?”

I will specifically address both of these questions in Chapter 3.

In this Chapter, I will look at the characteristics of the data without considering these questions. At this stage, I am more interested in drawing your attention to the characteristics of the data as they are.

A modest-sized collection of such foundry data can be plotted as a frequency histogram. This will show how many measurements are found in any given data range (represented by a bar of defined height in the histogram plot). The histograms of the foundry data show very skewed shapes that are left-truncated and no measurement values are plotted below a given value (either Limit of Detection or of Quantitation, or a Reporting Limit). An example is shown in Figure 2.1a and the histogram of the logarithms of the same data is shown in Figure 2.1b. Transforming the data by taking logarithms or with software that allows one to plot on a logarithmic scale is very useful with the display of skewed and left-truncated data because it helps to demonstrate this pattern better. Note that, in Figure 2.1b, the largest bar is at a $\log_{10}$ value of -3 which is the lower limit of detection for lead as measured in this type of monitoring. We will encounter a detailed discussion of statistical graphics and tabulations as ways of displaying data in Chapter 4.

In the past, the assumption has been made that the shape of the data set can be modelled as a lognormal distribution in which the histogram is skewed to the left and drawn out to the right (see Figure 2.2a). When we transform such data by taking logarithms, the histogram approximates in shape to the characteristic bell (see Figure 2.2b) of the so-called “Normal” or Gaussian distribution (see for example: Leidel N et al., 1977). Harvey and colleagues tried to model industrial hygiene data with a left-truncated or left-censored lognormal distribution and then tried to demonstrate a fitting process for some simulated data (Harvey RP et al., 1981). They then tried to apply this to some collections of measurements of occupational exposures to styrene vapour encountered in boat building. Having tried this approach myself, I find it inappropriate with the foundry data because the pattern of the sample data distribution is distorted by the presence of a large amount of data at the Limit of Detection and the Limit of Quantitation piled up like a “logjam”.

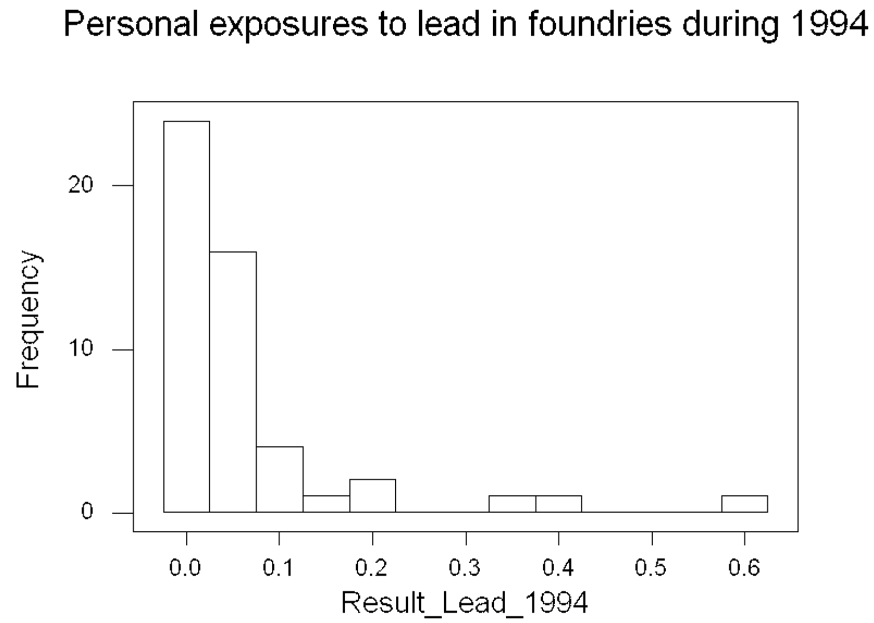


Figure 2.1a Histogram of personal exposures to lead in Foundries monitored by CTi during 1994.

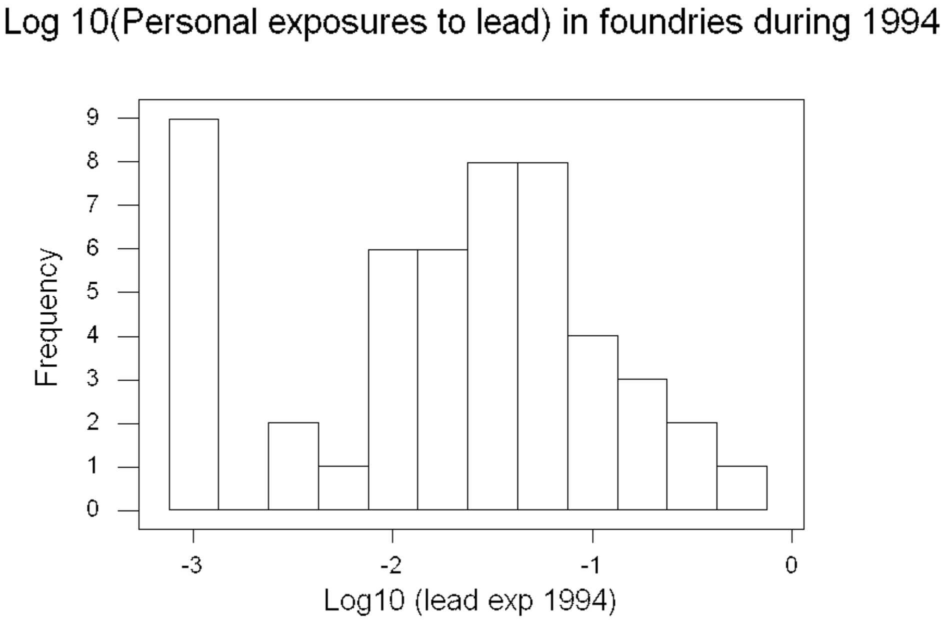


Figure 2.1b Histogram of the $\log_{10}$ values of personal exposures to lead in 1994 shown in Figure 2.1a

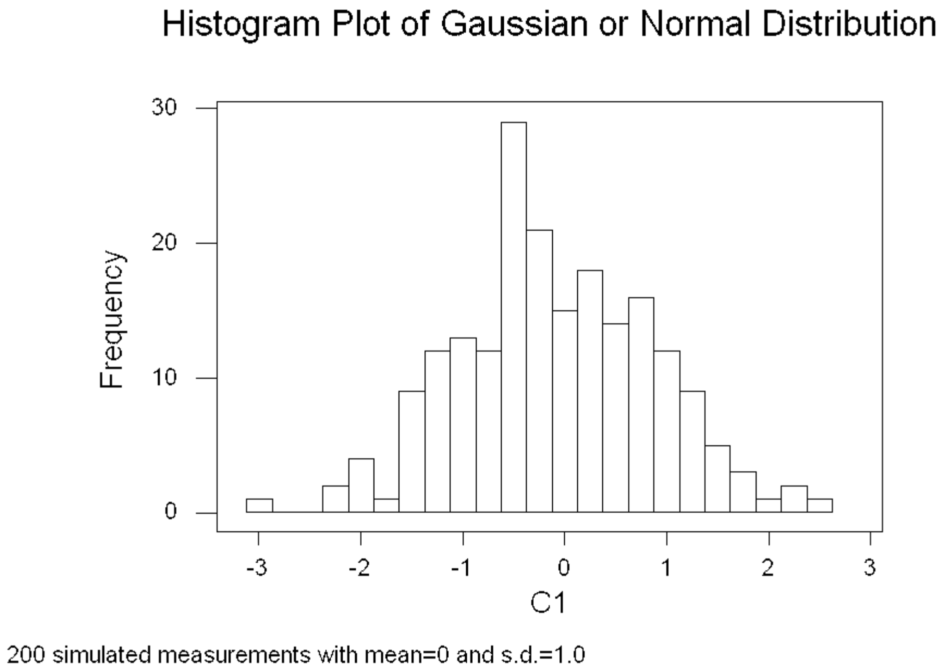


Figure 2.2a Histogram of 200 simulated measurements sampled from a Gaussian or Normal Distribution with a mean of zero and a standard deviation of one.

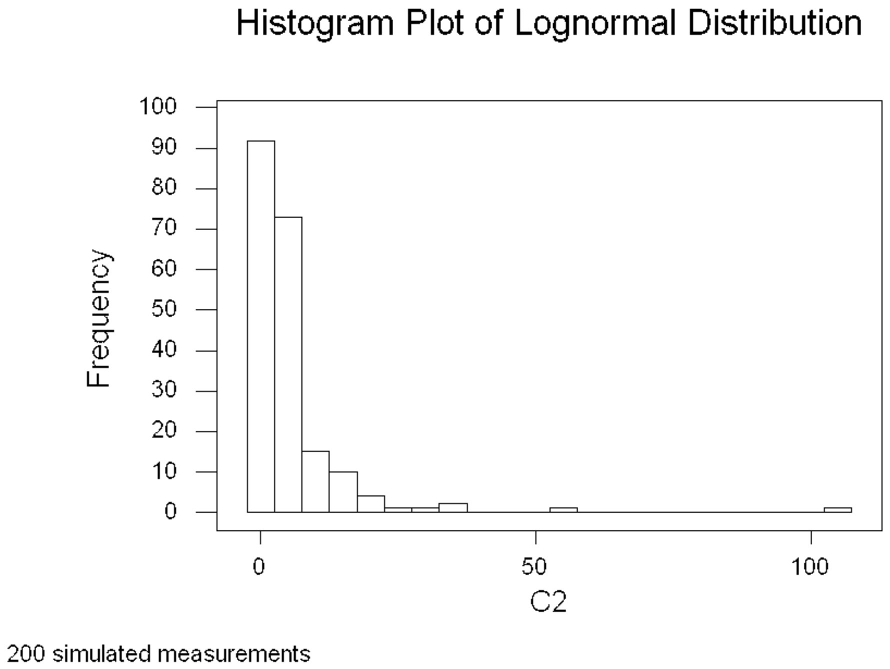


Figure 2.2b Histogram of 200 simulated measurements sampled from a lognormal distribution.

In modern statistical practice, many practitioners prefer not to impose a theoretical model on such data sets. Instead, we use either a distribution-free (or nonparametric) approach or what is now termed a data-driven or exploratory approach (see for example: Sprent P, 1998; Sprent P and Smeeton NC, 2001; Mosteller F and Tukey JW, 1977; Hoaglin DC et al., 1983, 1985), in each case allowing the data “to speak for itself”.

In this work reported here, I have used nonparametric, data-driven and exploratory approaches to reveal patterns in the data (see Thompson J M, 2000). This approach is important with the foundry data because not only do the shapes of the sample data sets appear to be “left-truncated” and skewed but also they appear elongated in the upper tail (i.e., in the range of the largest exposures).

Although the foundry database appears large, it does contain monitoring results for many different analytes, some of which appear as only small collections of observations with the focus being on dusts and gases of major interest. The real richness of the data set lies in its links with associated data that describes the monitoring context and properties of the data values (e.g., units of measurement, etc.). This associated data is known as metadata (i.e., data about data) and is very important in data analysis when investigating the effects of various factors on the values of the data being analysed.

Some data can be linked together to explore relationships between variables measured at the same time. For example, is there a consistent relationship between respirable silica and respirable dust across all the different processes, job types and foundry sizes and types? In other words, does the monitoring context matter? Some industrial hygienists have suggested that monitoring respirable dust could serve to act as a surrogate for monitoring respirable silica. I have had the opportunity to test this kind of hypothesis using the foundry exposure database and I illustrate the results of such analyses in Chapter 5.

An important problem arose when Castings Technology International queried the validity of some measurements of respirable dust and respirable silica and asked me to devise an approach to highlighting and then filtering out suspect data.

The presence of excessively large logjams at the Limit of Detection/Quantitation was an informative factor in this respect as was the link into a Quality Assurance process that enabled cross-checking and validation to be done for the period of time that was potentially suspect for some of this data. The suspect data was removed from the database so enabling the quality of data analysis and interpretation to be maintained at a high level.

In a previous collaboration with Damien McElvenny from the HSE Epidemiology Group and Diane Llewellyn an industrial hygienist from HSE Bootle, I developed techniques for examining small sized data sets from the HSE National Exposure Database (NEDB) (see Thompson, J. M. et al, 2003). In that research, I examined data sets varying in size from less than 10 measurement values to close to 100. One useful finding I made from that research was concerned with what can usefully be extracted from various small data sets and with what level of confidence. I evaluated various methods in that work for reliably estimating confidence levels.

From that research, I came to the realisation that, in using such data sets to inform decisions about Regulatory Limits, understanding the characteristics of the uppermost portions of the data was likely to be of greatest value.

I also began to realise that when evaluating Health and Safety Regulatory impacts on exposure patterns, the data pattern might well change shape by pushing the upper half of the data distribution towards the median but not necessarily affecting the median or the lower percentiles very much. In other words, changing the upper part of the sample distribution shape may have little effect on the sample median and the lower part of the sample distribution beneath the median. This is an extremely important consideration when attempting to devise meaningful and testable hypotheses relating to changes in distributional shape.

One of my specific goals in this project was to devise suitable, reliable approaches to evaluating such shape changes in the data sets, year on year. The purpose was to help in the evaluation of the potential impact of legislative and regulatory events, e.g., introduction of the Control of Substances Hazardous to Health Regulations (COSHH), changes in regulatory exposure limits, etc.

With my experience in evaluating and validating medical diagnostic and prognostic tests and equipment, I have come to the view that occupational exposure monitoring is actually a form of medical diagnostic and prognostic testing. It should therefore be subject to similar kinds of investigation in terms of evaluating its efficacy, reliability and accuracy for prognostic activities relating to predicting future health morbidity linked to that exposure pattern.

Similarly, such databases have immense potential when retrospectively evaluating occupational exposure records to aid diagnosing causes of current morbidity and predicting the subsequent changes in morbidity. This would be far more satisfactory than using conventional epidemiological approaches to the assessment of impacts of exposure linked to mortality.

Morbidity linked to occupational exposure patterns, year by year, has important relevance to Quality of Life and thus has a possible impact socially and economically as well as on the productivity of those occupationally exposed to hazardous substances.

CHAPTER THREE
SAMPLE COLLECTION AND ASSAY, DATA COLLECTION, RETRIEVAL, PREPARATION AND VALIDATION

Sampling Strategy in Occupational Exposure Monitoring

Strategy concerning where, when, what, on whom, how often and for what purpose monitoring is done is most often empirically determined by the hygienist's judgement guided by his or her experience, skill and knowledge. Judgemental sampling of that kind has a role and a justification from the need to monitor those people most at risk and those areas and processes most likely to present a risk. This approach disturbs statisticians who are unfamiliar with needs of hygienists in their evaluation of occupational exposure risks.

Conventional Representative Sampling

The reason for the concern from those involved in applying conventional statistical sampling strategy design is that the sampling appears excessively biased towards those at greatest risk and is likely therefore to be unrepresentative of the general population of employees exposed to the occupational hazard being monitored. In that sense, it may not adequately be able to assess the risk of all employees but of course that is not the point of such monitoring and it is not particularly intended to be random sampling of the kind used to assess the overall risk of all exposed employees.

Appropriate Sampling

Leidel, Busch and Lynch (1977) in Chapter 3 of the US NIOSH "Occupational Exposure Sampling Strategy Manual" offered an approach to what is usefully termed "appropriate sampling" of the "maximum risk employee" (Leidel N et al., 1977). Additionally, it offered guidance on strategies for monitoring other employees at lesser risk but requiring adequate risk assessment in a cost beneficial and practicable manner that impacts on employers without prejudicing their business survival. It has thus provided valuable guidance that has served industrial hygienists well in the assessment of occupational exposure risks.

Developments in statistical sampling strategy design since then have opened up new opportunities to do effective and reliable sampling that is focussed on the need to assess those employees most at risk whilst still gaining valid information about the spread of exposures experienced by all other employees at risk.

The area of adaptive statistical sampling strategy design (see, e.g., Thompson SB and Seber GA, 1996) offers this particularly in the form of ranked set sampling strategy design (see, e.g., Chen Z et al., 2004).

The approach adopted in gathering the occupational exposure monitoring data for the foundry industry members and its other clients by Castings Technology International and its predecessors, over the past 30 years or so, has typically involved focussed sampling on employees at maximal risk (Wells D and Greenall A, 2004). In addition, area sampling was undertaken in most process and work areas and general sampling of employees at lesser risk was done to obtain a general picture of the overall exposure patterns.

The patterns of exposure suggest that they have generally succeeded rather well in capturing both kinds of exposure data in a way that appears reasonably representative although the patterns will inevitably be somewhat biased in favour of those employees in the higher risk groups. The database that has been prepared for HSE NEDB by Castings Technology International is thus a very valuable resource for a wide variety of investigations beyond those reported here, as well as in the parallel and complimentary report to the HSE prepared by Castings Technology International.

Assays of Exposure Samples:
Issues of Validation, Quality Assurance and Reporting

From the viewpoint of the quality of measurement, both physical sampling processes and methods of physical sample preparation and subsequent analysis are of fundamental importance. Many potential problems can be identified by careful examination of original records from the surveys undertaken, as well as from the receiving analytical laboratories.

In discussions with staff at Castings Technology International about methods of measurement of total and respirable silica dusts, it became apparent that there were considered to be problems with a method deployed in one of the analytical laboratories used (Wells D and Greenall A, 2004). That method appeared to seriously underestimate the silica content of dust samples. When staff at Castings Technology International became aware of this, method comparisons were undertaken to demonstrate the problem and this was then followed by implementation of a much more rigorous analytical validation and quality management system, thereby eliminating that source of bias.

I was requested to evaluate the reliability of the data from the period in which there appeared to be this source of analytical bias. A key symptom of this problem of serious underestimation bias was the very large proportion of “non-detects” (measurements at or below the limit of detection/quantitation for silica dust). Having identified the suspect period of data collection, I recommended the removal of that unreliable data from the database. This appears to have been the only major problem in chemical analysis but it is a useful and salutary lesson to bring to the reader’s attention.

In checking the validity of data subsequently reported in various ways and for various purposes, it must be recognised that there will be occasional mistakes made in physical sampling, sample preparation and analysis.

In addition, there will inevitably be transcription errors in the transfer of data and associated metadata from one form of record to another. In the case of the foundry exposure data, all records were transcribed from paper-based monitoring reports and laboratory records into a password-protected Microsoft (MS) Excel Spreadsheet File laid out to allow it to be ultimately imported into the HSE NEDB.

Staff at Castings Technology International undertook that transcription process within a system of rigorous auditing and editorial quality management. Despite this, during my initial auditing of data passed to me for investigation, I found further errors. I must emphasise that this does not reflect at all adversely on the quality management of the original transcription process but merely serves to demonstrate the need to perform several stages of such checking processes. I used a number of exploratory data analysis methods as well as tools within Excel in conducting this transcription audit, as well as standard technical editorial techniques.

Data Retrieval

Castings Technology International presented me with data in the form of password-protected MS Excel files, at various stages in the building of the database so as to enable me to develop the statistical methodologies in parallel with database compilation. I used standard tools available in MS Access to extract an unprotected version of that MS Excel file. I then used the Advanced Filter within Excel to extract subsets of data and associated metadata into smaller and more specific Excel files.

Additionally, Andrew Greenall of Castings Technology International used Excel’s Pivot Table method to extract, into separate MS Excel files, data and associated metadata that was linked together in specific ways (especially in relationship to simultaneous measurement) to allow relationships between specific groups of measurement variables to be explored. An example of that kind of linking of special interest was for total dust, respirable dust and respirable silica.

From these various smaller files I could then extract relevant data and metadata into the Minitab statistical software package (version 13) that I used for development of appropriate methods of statistical analysis and for subsequent data analysis.

Data Preparation

Much of the prior data preparation for statistical analyses within the Minitab environment involved stacking and unstacking of data columns, selection of subsets of data and metadata, sorting and ranking, etc. in order to get sets of data and metadata into appropriate formats for analysis and plotting. Some data preparation also involved early stage mathematical transformations and also undertaking initial exploratory statistical analyses as a prelude to using the outputs from those as inputs to other analyses (e.g., undertaking Letter Value Analysis to provide Letter Values as inputs to Lowess plots or analysis of variance, etc.).

Data Reporting

A problem that I have previously encountered when dealing with data on chemical and physicochemical analysis undertaken by a number of laboratories relates to inconsistencies in reporting that data, even though the analyses may have been subject to quality management of a good standard. Each laboratory may have its own consistent reporting practice, sometimes enforced by the United Kingdom Accreditation Service (UKAS) if the laboratory is so registered or is a member of a Proficiency Scheme.

This can obviously cause problems when different laboratories reporting to the same contracting organisation do so without reporting practice being standardised and managed for consistency by that organisation.

CHAPTER FOUR
STATISTICAL METHODOLOGY, INCLUDING STATISTICAL GRAPHICS AND TABULATIONS USED, WITH ILLUSTRATIONS OF APPLICATIONS USING FOUNDRY INDUSTRY DATA

How is the Foundry Data explored, evaluated and analysed?

Because the foundry data does not fit any statistical data distribution model, e.g., truncated (censored) lognormal, for a variety of reasons, I have approached the exploration and analysis of the data using methods of exploratory, robust and resistant statistics, including nonparametric or distribution-free statistics (see below). This approach protects us against the problems of more classical, parametric statistics that generally relies on the data behaviour conforming to that which you expect from a specific model.

Classical, parametric statistical methods are vulnerable to potentially serious misinterpretation when used with data such as that from the foundry industry. This is because there is no straightforward way in which the data could be reliably transformed mathematically to fit approximately to a Gaussian distribution, and also because of the presence of extreme or outlier data to which such methods have zero resistance (see, e.g.: Goodall in Hoaglin DC et al., 1983).

There is insufficient data in each category within the foundry data collection for any appropriate and satisfactory way of demonstrating even a very approximate fit to any formal data distribution model, such attempts should be regarded as futile and unsafe.

The exploratory approach that I have adopted enables me to characterise the key features of a data set that show its shape in a systematic way that facilitates my development of testable hypotheses on trends in data patterns potentially linkable to regulatory and legislative events. In this respect, a particular interest was concerning the possible influence of foundry size on the lag between when such an event occurred and its possible impact on observed occupational exposure patterns.

In addition, I have adapted and modified a number of statistical methods to refine this exploratory analysis to reveal in a rigorous and robust manner underlying trends not otherwise so readily shown by standard approaches. These methods enable visual demonstration of trends, if present, in a way that statistically unsophisticated readers can reliably understand and interpret.

Issues relating to Hypothesis Testing

A major issue that I debated in discussions prior to establishing the project on which I am reporting here, was that concerning whether all testable hypotheses should be formulated and, hence also approaches to their being tested, at the beginning of the project or even in its definition and planning. As an applied statistician with experience in environmental data analysis, I was very wary of such an approach because it necessarily presupposes that testable hypotheses can be properly formulated before you are really aware of the nature of the data and metadata that are being catalogued. Nor can you fully appreciate the data structure until it has been adequately explored.

An attempt had been made in discussions, between CTi and HSE staff and me, to draft such a set of hypotheses but, quite quickly into the early phases of compiling the foundry exposures database, it became apparent that many of those hypotheses would not be testable.
Later into the project, it was found that other hypotheses were also not testable. So the initial exploratory analyses became a major influence on the formulation of meaningfully testable hypotheses.

Exploratory Data Analysis

The initial approach to exploration that I used involves techniques that provide tabular and graphical summaries of key features of the data which begin by sorting the data into order of increasing value and then ranking those sorted data values. Ranking is a powerful method of preserving important features of the data but protecting the data analyst from the distortions of interpretation to estimates of the "average" and "spread" of a data set caused by the presence of extreme or outlier data values. Two important techniques that I used in the early stages are briefly described below:

(a) Letter Value Analysis

This technique, introduced by John W. Tukey (see Mosteller F and Tukey JW, 1977), provides summaries starting with the median and working out away from that in both directions through the ordered data set.In so doing, it gives a progressively more detailed view of the pattern of spread of the data towards the upper and lower extremes (the minimum and maximum).

The term Letter Value relates to the tags given by Tukey to the summary values estimated. The first summary value, the median, is given the tag M. The next Letter Values away from the median are the Upper and Lower Fourths (tagged F_U and F_L) or Hinges (tagged H_U and H_L) corresponding approximately to quartiles with 25% of the ordered data respectively above and below them. Then moving further out, we have the Upper and Lower Eighths (tagged E_U and E_L) with 12½% respectively above and below them in the ordered data set and so on outwards to the upper and lower extremes.

Each step away from the median involves halving the percentage respectively above and below the corresponding upper and lower Letter Values giving a progressively more detailed look at the so-called "tails" of the ordered data set.

Included in this analysis as a tabular display are the Letter Value Spreads, the differences between the upper and lower values, e.g., the Fourth Spread (which corresponds approximately to the Interquartile Range). This encompasses the middle 50% of the ordered data and is a very useful spread measure. Other Letter Value Spreads also give valuable information about the data set pattern.

Also included in the tabular display are the averages of each pair of Letter Values (Upper and Lower). These are the so-called "Mid" or "Mid-summary" values and the trend of these values down the Letter Value Display gives valuable information about the shape of the data set. If the values of all the "Mids" stay roughly the same, then the data set is roughly symmetrical. However, if the "Mid" values either increase or decrease systematically down the display, then we can see that the data set is skewed either to the left or right and the extent to which skewness is present.

A typical Letter Value Display for the Foundry data is shown in Table 4.1. If the Letter Values for an analyte (e.g., respirable silica) are plotted as a function of year then we can explore trends in the pattern within a data set over time. From such a display, I began to formulate testable hypotheses about data trends.

Table 4.1

Letter Value Analysis Table for Personal Exposures to Respirable Silica in mg m^{-3} (CTi Alvechurch data) for 1990

	Depth	Lower LV		Upper LV	Mid	Spread
N = 118						
M	59.5		0.050		0.050	
F	30	0.030		0.084	0.057	0.054
E	15.5	0.021		0.207	0.114	0.187
D	8.0	0.013		0.322	0.168	0.309
C	4.5	0.010		0.355	0.183	0.345
B	2.5	0.007		0.750	0.379	0.743
A	1.5	0.004		1.124	0.564	1.121
*Extr	1	0.003		1.208	0.605	1.205

* Extreme values, minimum and maximum

For example, I was interested to test whether Upper Letter Values were influenced by introduction of new legislation or by downward changes in regulatory exposure limits. I also wanted to investigate whether it was the Upper Letter Values only that were influenced by such events, in other words the occupational data set was changing shape but the “average” (median) exposure did not change. In which case, if the exposure limit was set appropriately and controls on exposure were well implemented, fewer foundry workers would be put at risk of high occupational exposures.

So the Letter Values and their spreads and mid-summaries that I estimated could then serve as inputs to other more sophisticated analyses including those involving testable hypotheses.

(b) Box and Whisker Plots

These graphical displays are valuable for showing key features of the shape of a data set. The upper and lower fourths forming the upper and lower sides of the box and having a line within the box between these sides indicating the position of the median on the measurement scale. The so-called “whiskers” show features of the spread of data beyond the middle 50% that could not be could not be considered unusual.

Data values beyond the boundaries of the “whiskers” are plotted individually so as to draw them to our attention in order to consider whether they might be deemed extreme (outliers). An example of a box and whisker plot for some data from the Foundry exposure database is shown in Figure 4.1 and illustrates these features.

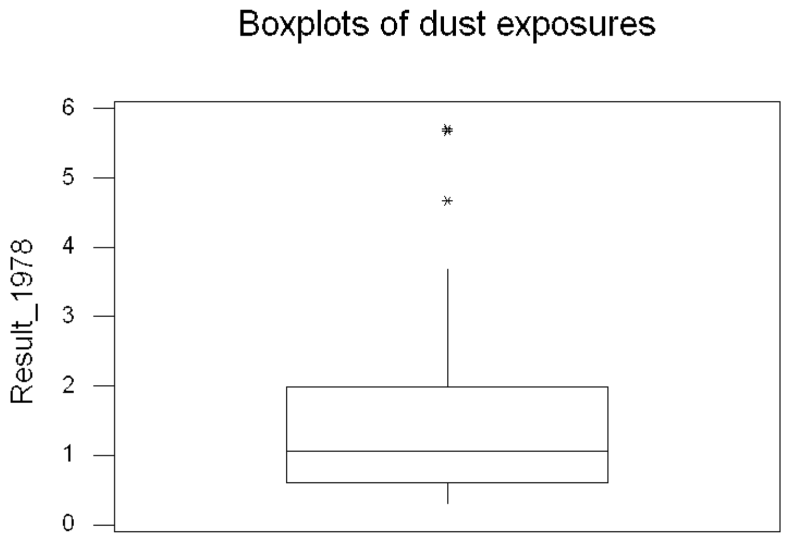


Figure 4.1 Box and whisker plot of dust exposures. The line within the box is the median and is located closer to the bottom of the box indicating skewness of the middle 50% of the data. Two outlier data points can be seen above the upper whisker showing extreme exposures.

When comparing data sets, for example, year by year, it is useful to plot boxplots for each year side-by-side for useful comparisons (see Figure 4.2).

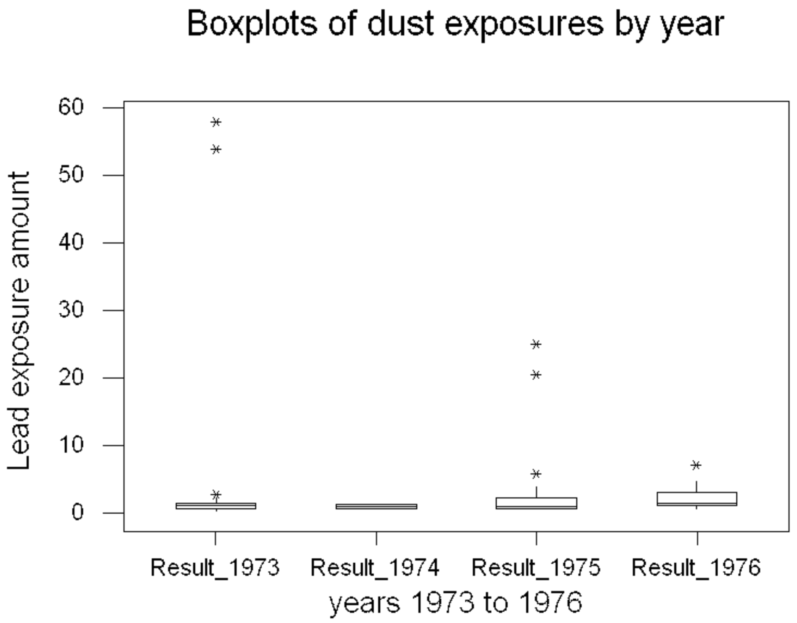


Figure 4.2 Side-by-side boxplots allowing year by year comparisons. Note, however that the outlying (extreme points) are spread over a wide range of values and so the vertical scale is not very appropriate for showing the detail of the boxplots.

Within the boxplot, a confidence interval on the median can be included to enable comparisons to be made of the similarity of medians in the manner of a graphical one-way Analysis of Variance (see Figure 4.3).

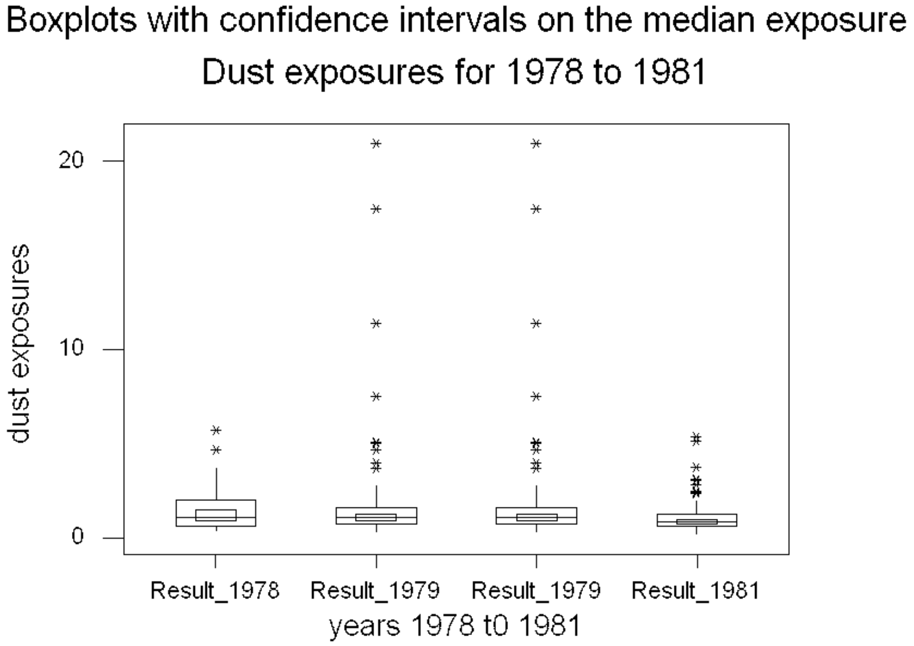


Figure 4.3 Side-by-side boxplots with inner boxes showing the 95% confidence intervals on the medians. The confidence intervals were calculated in Minitab 13, using the Sign Test.

If the confidence interval boxes within the boxplots overlap well, this indicates that the medians are not significantly different. If such confidence interval boxes do not overlap and are reasonably far apart then the medians can be considered as probably significantly different.

If the overlap or gap between confidence interval boxes is only small then the conclusion from the graphical comparison is potentially ambiguous and suggests the need for a more formal multiple comparisons test that is set up to avoid such false or potentially ambiguous conclusions.

Refining Hypotheses following the Initial Exploratory Data Analyses

Whilst there is potential for biased hypothesis setting by formulating hypotheses after such initial exploratory analyses as described above, if the process of formulation is carefully done then the risk of bias is unlikely to be a serious problem.

This problem is discussed in detail by Diaconis (Chapter 1 in Hoaglin DC et al., 1985) with various remedies offered including:

(a) quantifying multiple comparisons, and hence reducing the risk of ambiguity, using a simple system of inequalities to set safe thresholds for significance for the alternative hypothesis to avoid false interpretations (this is a standard and well-known method),

(b) trying out the hypothesis on fresh data (this is distinctly possible with the foundry exposures data because initial exploration on a small sample from the data is possible, readily allowing this option),

(c) borrowing strength from more-or-less parallel situations (again, this is possible with the large size of the foundry exposures data set),
(d) cross-validation by trying out the exploration on a relatively small random sample of the data and then testing the new hypothesis on the remainder of that data sample (again, this is a well-established approach),
(e) the exploratory analysis can be repeatedly tested using the technique developed by Efron (see Efron B and Tibshirani RJ, 1993) known as bootstrapping and/or by the technique known as jack-knifing developed Tukey (see, e.g., Mosteller F and Tukey JW, 1977).

Deploying one or more of such "cross-checking" techniques can give us confidence that conclusions drawn and interpretations made based on hypotheses formulated after the initial exploratory analyses can have strong validity. In particular, I have used methods (a), (b) and (c), described above, to give that strength to hypothesis testing.

In the field of process monitoring, Box G and Luceño A(1997) argue the case for using control charts playing an inductive role, enabling the data analyst to suggest unexpected hypotheses for consideration. They believe that hypothesis generation, as with any learning process, is very important: "we cannot test an hypothesis we don't yet have".

Robust and Resistant Nonparametric Data Analysis:

What is resistance? Data analysis methods that possess resistance protect us against the untoward influence of outliers as a result of insensitivity to localised misbehaviour of some of the data compared with the bulk.

What is robustness? This may be distinguished from resistance in that it implies "an insensitivity to departures from an underlying probabilistic model" (Hoaglin DC et al., 1983), e.g., departures from a truncated lognormal model.

The techniques that I have used and/or modified in this research project possess these useful properties.

Nonparametric Data Analysis using Ranking Techniques

These techniques enable us to evaluate when trends in occupational exposures may have occurred, over the many years that monitoring has been undertaken. Using ranks protects us against excessive influence of extreme values and offers sound methods that are distribution-free.

In particular, I have used the grouping of data on an annual or multi-annual basis to evaluate whether significant changes have taken place using rank-based one way analysis of variance, followed by multiple comparisons to determine when significant changes may have occurred. This way, I could identify changes that might be described approximately as stepped as opposed to gradual. Different techniques need to be used to differentiate step change from gradual change, so the limitations of each approach need to be recognised in the formulation of hypotheses.

Kruskal Wallis One Way Analysis of Variance (ANOVA) by Ranks

I used this method with year or year group as the Factor to be tested for its effects on the median of exposures to a given dust or gas. The Kruskal Wallis ANOVA involves ranking all the data, in the set to be analysed, across all the groups and allocating the ranks back to the groups to which the original measurements belonged. Then the average ranks of the groups are compared (see, e.g., Gibbons JD, 1976). If the test statistic calculated in this initial comparison is sufficiently large it will suggest that there may be one or more groups with a significant difference to the remaining groups but it will not allow us by itself to judge which might be different. The Kruskal Wallis test is well recognised as a sound test that is highly resistant and robust.

To discover whether a group or groups might be different from others included in the ANOVA, we need to perform a “multiple comparisons” test.
I applied this test firstly to data grouped by year and also as three-year groups, using all the relevant data for an analyte for the relevant exposure context. This enables comparison of median exposures between the groups.

However, I was also particularly interested in modifying the Kruskal Wallis test to evaluate trends with upper 50% of the data to determine whether there was a trend in Upper Fourths (F_U) and also the upper 25% of the data to examine trends in the Upper Eighths (E_U).
Both such evaluations should be useful in examining groups of workers most at risk.

If data sets were large enough, I used the same approach to examine the behaviour across the groups with data greater than the Upper Eighths to test for trends in the Upper Sixteenths or Upper D (D_U).

Multiple Comparisons following Kruskal Wallis ANOVA

The purpose of “multiple comparisons” tests is the determination of which groups differ significantly from others. Various “multiple comparisons” tests are available which protect against false conclusions by setting high thresholds for significant differences between all possible pairs. I used O. J. Dunn’s multiple comparisons test for such comparisons in some such cases (see Gibbons JD, 1976).

As an alternative to pursuing this approach, I also chose to examine trends using Lowess (see below) and Cumulative Sum techniques (see below) with Group Medians, Fourths, Eighths, etc. The reason for doing so was connected with patterns observed in examining simple plots of these summaries versus the one-year groups or 3-year groups. In such plots changes appeared not necessarily to show straightforward trends with time perhaps indicating the additional influence of other factors, besides just time (e.g., changes in processes undertaken in certain foundries).

Lowess and Cumulative Sum plots are capable of revealing more subtle trends than just step changes and I considered this to be a more useful approach in this context.

Robust and Resistant Regression

I could assume that a trend versus time might be crudely considered to be linear with some function of time (e.g., $y = a + bt$ or $y = a + bt + ct^2$). It might then be a robust and resistant regression might be useful but that assumes that a straightforward function could serve as a model for the trend. Although I could find some crude fits this way, they were not informative in relationship to possible legislative and regulatory events occurring along the time-line. So I spent little effort exploring this approach.

Another valuable exploratory technique: Locally Weighted Regression for Smoothing of Time-varying Trends (Lowess or Loess)

If a trend is not likely to be a straightforward function versus time then it is better to pursue an approach based on finding a locally smooth function of the variation with time. Cleveland in 1979 devised a “locally-weighted” regression-type of smoothing technique (see, e.g.: Cleveland WS, 1994 and Bowman AW and Azzalini A, 1997).

In the version of Minitab that I used to explore the application of this technique (v13), I was able to check the degree of fit of the Lowess, with the Lowess parameters at various settings, to optimise the model of variation with time.

I had recommended, to staff at Castings Technology International, that they obtain Minitab, for some exploratory data analysis they wished to undertake. However, the version they obtained (v14) very regrettably does not appear to include features in the lowess algorithm to allow the user to check how good a fit the lowess achieves. Being able to check the goodness of fit is very important in any attempt to reliably explore links between events along the time line and changes observed by the Lowess smoother technique.

This technique has been of great value in this project for exploring trends but I cannot recommend Minitab versions later than version 13 for this purpose at this point in time. Other software such as S-plus would now be much more appropriate.

Adapting Lowess for Analysis of Trends of the upper portions of data sets over Time: do the data distributions change shape over time?

Usually, Lowess is used on the whole data set in each group along the time line and it is very robust and resistant as a technique in this respect. Applied in this way, it is somewhat like smoothing a plot of the group medians.

The modification of the application of the Kruskal Wallis ANOVA, which I applied the technique to the upper 50%, the upper 25% and the upper 12½% of sets of data, provides useful insights. Likewise, this approach with Lowess enables us to observe trends in the portions of data of most interest especially in terms of the impact of trends for the most exposed groups of workers. Examples of such analyses are shown in Figures 4.4 to 4.7.

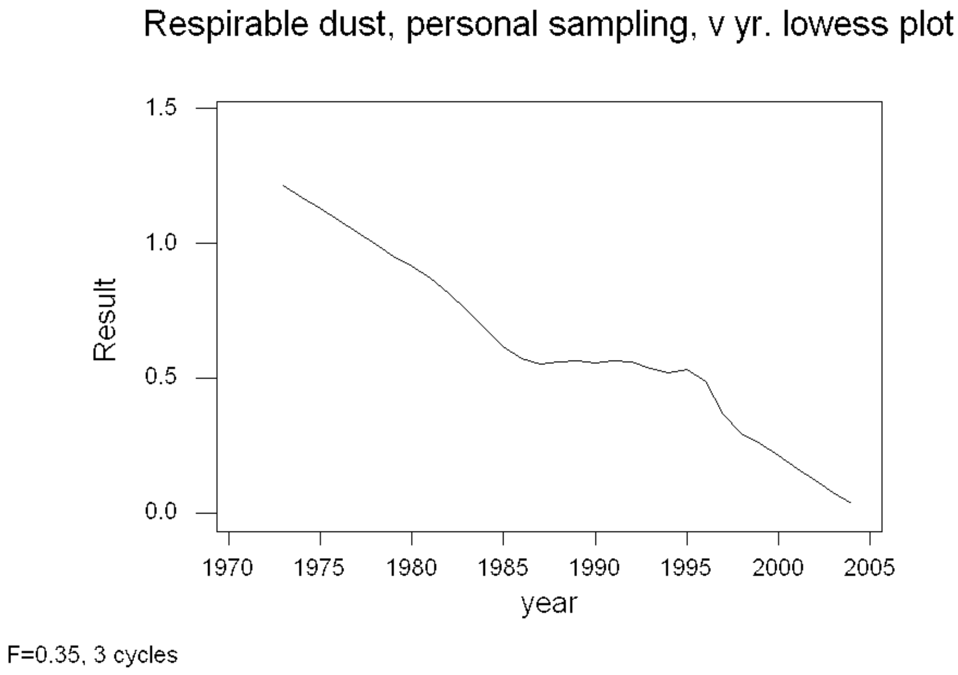


Figure 4.4 Lowess plot of the annual respirable dust personal exposures collected by CTi. The actual data within each annual collection of exposures is omitted for clarity.

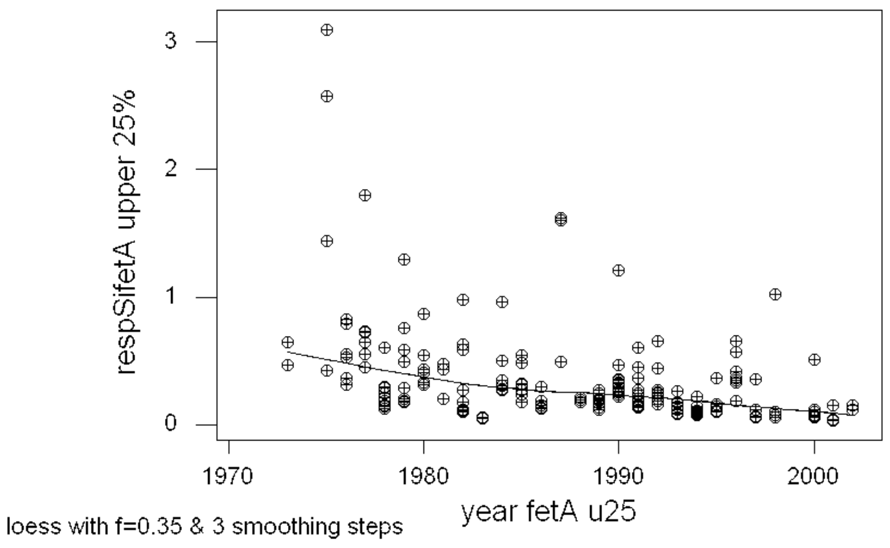
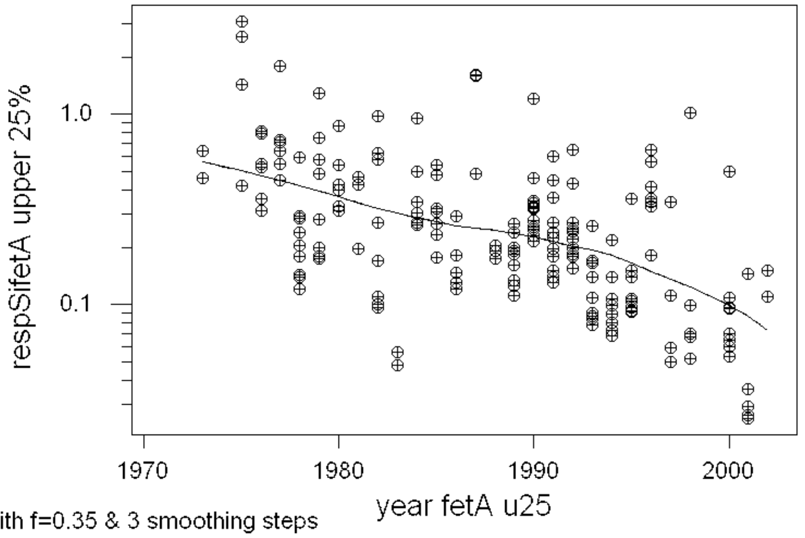


Figure 4.5 Lowess plot of the upper 25% of annual personal exposures to respirable silica during fettling processes collected by the Alvechurch branch of CTi. Here the actual exposure data is included but it should be noted that data points of the same magnitude may be overplotted and thus not distinguished as distinct points.



loess with f=0.35 & 3 smoothing steps
respSi & loess on log scale

Figure 4.6 Lowess plot of the same data as in the previous plot but with the respirable silica scale in $\log_{10}$ format

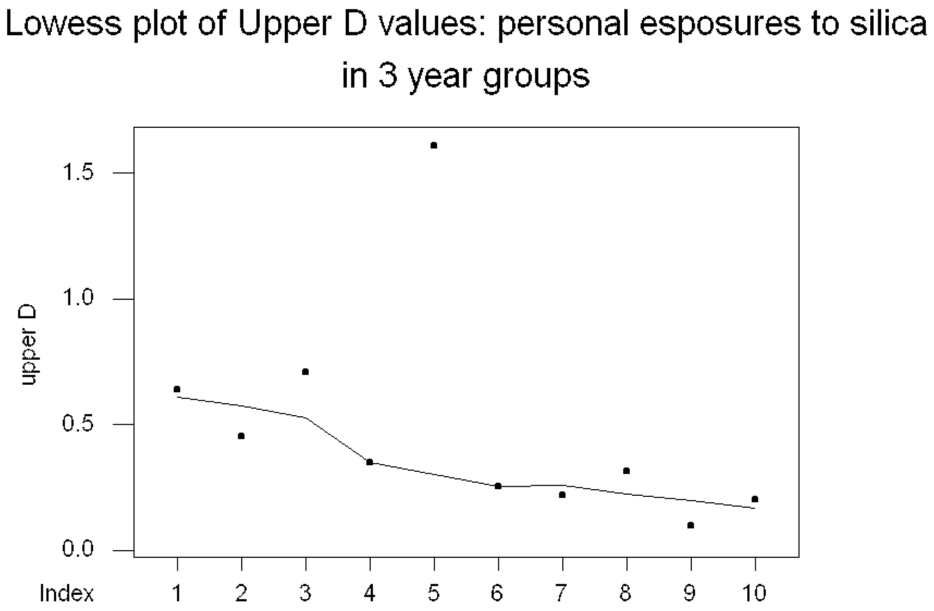

Lowess plot of Upper D values: personal esposures to silica in 3 year groups

Figure 4.7 Lowess plot of the upper D values (upper sixteenths) of personal exposures to respirable silica in three-year grouped data collections covering the period from 1973 to 2002.

Adapting specialist techniques from Statistical Process Control: Using Cumulative Sum Techniques

The Cumulative Sum techniques are ones in which successive deviations from a target mean for a process are accumulated and are then plotted in the form of control charts known as cumulative sum or Cusum charts. Such charts are often plotted as two "one-sided" charts, so that upward deviations are plotted separately from downward deviations and are often shown together on the same display (see Box G and Luceño A, 1997).

This approach is one that enables the detection of small persistent shifts from the target mean, either upwards or downwards, that are not otherwise easy to detect. We can use these plots to identify the period during which the shifts started to become significant, so enabling us to target investigations of possible causes. When significant deviations, beyond set thresholds, are found the chart is generally reset so that any further changes can be detected. That approach is very powerful for process control.

We may have an interest in reviewing and exploring periods of occupational exposures retrospectively, for a process or job function. We may wish to examine whether there are shifts that are either up or down in some parameter of the distribution of observed exposures, but do not have a particular target in mind. Instead of using a target mean, we look at the average of the parameter of interest over the "span" of the review period (from its start to the end of that period) that we wish to investigate. We can then use retrospective cusum charts, also known as cusum span charts, as described below, which help us to identify adjacent periods of time, in which the values of that parameter (e.g., conventionally, a process mean) differed significantly from one period to another. We can then compare these changes in the average value of that parameter with a timeline of occurrence of significant events that might impact on that parameter (e.g., introduction of the Health and Safety at Work, etc., Act, or the COSHH Regulations or changes in regulatory exposure limits). In this way, the method can provide valuable insights.

Cumulative Sum Span techniques

The long term cusum span chart is especially useful in detecting and assessing long term changes that would not necessarily be large enough to detect using the usual prospective control chart methods, including the conventional cusum chart (Taylor AL et al., 2002). This approach was originally described by Woodward and Goldsmith (1964) and incorporated into BS 5703 Part 2; this standard was issued in 4 parts over the period 1980 - 1982. The original BS 5703 standard on the use of cusum charts has recently been updated, again in 4 parts (2003). Williamson RJ (1985) developed alternative approaches to estimating short-term variability for use with cusum span charts. Taylor et al (2002) have recently published an approach to automated detection of shifts with cusum span charts.

A useful feature of this approach is the ability to perform significance tests to potential changes identified using the span chart to avoid over-interpretation of changes observed in the chart. A further valuable feature of the cusum span chart method is the ability to identify when a shift occurred and the size of the shift. I have modified this approach by applying it to the Upper Letter Values in order to evaluate when shifts of reasonable significance have occurred and to what extent. This is thus approaching the problem of the detection of shifts in a somewhat different manner to the Kruskal Wallis ANOVA method and may not necessarily give the same results because the discriminating capabilities and sensitivities of the methods will be different. No comparison has previously been made of these two approaches.

An illustration of the use of the cusum span chart applied to various Upper Letter Values is shown in Figure 4.8. This shows a major change in the mean annual value of the Upper D values for respirable silica. Above the upper D value, there is the uppermost 6.25% of the personal exposures. This change occurred in 1985. Prior to this, the mean annual upper D value was 0.692 mg m^{-3} and from 1985 the mean was 0.211 mg m^{-3}, about a 70% reduction and a dramatic impact on the average risk of the most exposed foundry workers. It is possible to dissect the shorter period from 1985 to 2001 to identify other possible changes within that period.

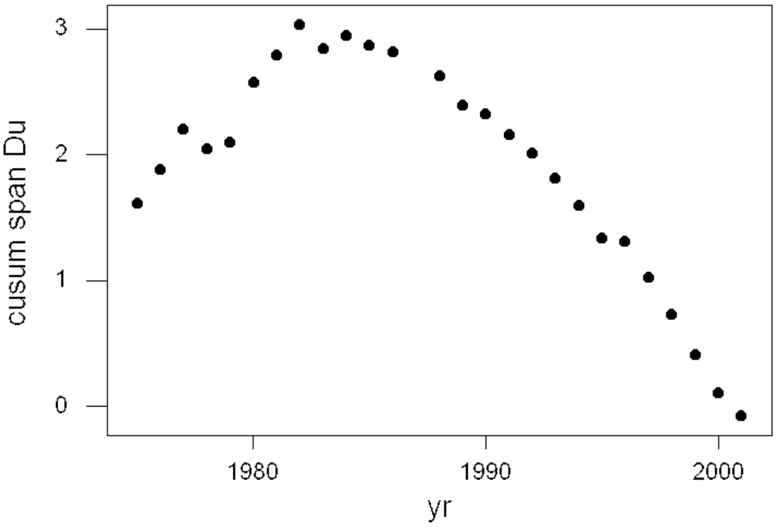


Figure 4.8 Cusum span plot showing a significant change in mean value of the Upper D values of respirable silica over 1975 to 2001, omitting the year 1987 when there was an exceptionally high value. The plot shows a change of direction around 1984, which indicates a reduction in the mean annual level of the Upper D values.

CHAPTER FIVE
TURNING DATA INTO INFORMATION: SOME EXAMPLES OF INTERPRETING DATA EXPLORATION AND ANALYSIS

Testing some Specific Types of Hypotheses illustrated with examples from the CTi Foundry Exposures Database.

Is there a relationship between respirable silica and respirable dust?

An interesting hypothesis posed by the CTi staff was whether measuring respirable dust could be useful as a surrogate measurement of respirable silica. This was tested by examining the sets of measurements in which there had been assays of total dust as well as of both respirable dust and respirable silica on the same physical sample. When looking at these sets of measurements across the complete database, no relationship could be discerned between the two respirable analytes and none existed with total dust either.

Another related hypothesis was: “Does a relationship exist between respirable silica and respirable dust for a particular class of foundry size?” From the lowess plot shown below, (Figure 5.1) this does not look to be a very promising idea. A further hypothesis of this type might be more context-specific and involve testing the relationship for a particular process in a particular class size of foundry but that seems not to be very sensible as an approach. If it is so context specific, then it is futile to pursue.

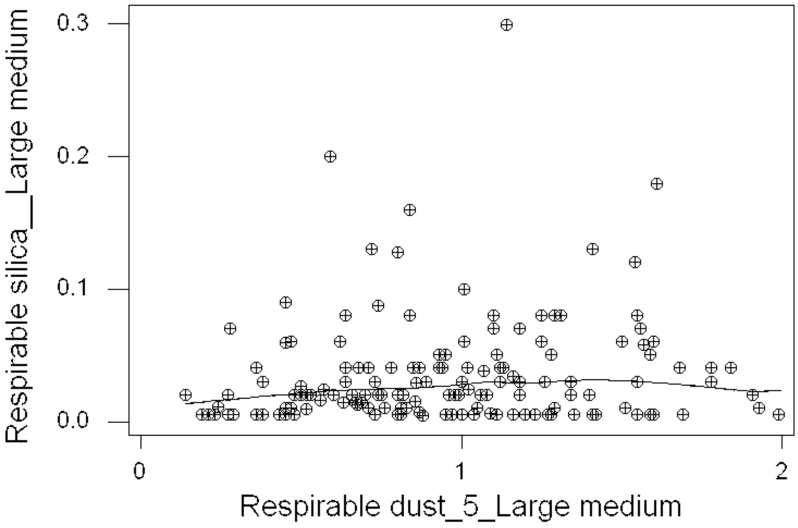


Figure 5.1 Lowess plot of respirable silica versus respirable dust (personal sampling) for large-medium foundries

A general class of hypotheses concerning whether there are changes in occupational exposure levels as a function of time.

Hypotheses of this type are especially interesting because they are likely to involve testing whether any changes might be observed that could be linked to:

- Changes in legislation (e.g., introduction of the Health and Safety at Work, etc. Act or Statutory Instruments) or
- Changes in regulatory occupational exposure limits for different hazardous substances

and understanding

- That such a possible kind of link might perhaps occur some time after the change of law was introduced and the lag might vary according to the size of foundry.

Within that class of hypotheses, we could formulate others that test whether possible observed changes might be different for:

- Different sizes of foundries
- Different types of foundries
- Different processes within foundries
- Different jobs within foundries
- etc.

I also posed the question earlier in this research report about whether the sample distributions of occupational exposures might change shape with time and that such change might be observed as major reductions in the upper levels of exposure with time. I expected that it could well be possible that there might be a much smaller effect on the median exposures and perhaps no effect on exposures less than the median.

In this chapter, I introduce some examples of such hypothesis testing in which I have used some of the techniques that I introduced in the previous chapter.

One of the most interesting techniques is the Lowess Plot. This Locally Weighted Regression essentially smooths a fit to the data for relationship that is probably not linearly a function of time. Unlike some smoothing algorithms that work only on a single measurement or measurement summary parameter at each time point, Lowess is similar to a group of algorithms known as kernel smoothers (see Bowman AW and Azzalini A, 1997).

In the first set of illustrations of analyses that I performed, I have shown how different groups of process within foundries have varied in overall occupational exposures to respirable silica. The two groups of processes are different in a number of respects. Fettling processes are inherently much dustier and it is likely to be harder to control exposures in such processes than many others in foundries, such as the group encompassing melting, mould production and core and mould production (mmpcmp). This is illustrated in Figures 5.2 and 5.3, respectively. The lowess lines in these plots correspond to some kind of weighted moving average of the measurements in any given year.

Clearly, the fettling exposures are greater than the other group but they reduce significantly more obviously year on year as the introduction of the Health and Safety at Work, etc. Act approaches. Then there seems to be period in which no further reduction occurs for some years followed by a further period of reducing exposures. Little change seems apparent for the other group (mmpcmp) over the years but this is probably because it is genuinely a cleaner group of processes requiring less investment in exposure reduction than fettling in order to comply with legislation.

This approach using the complete annual sets of data for each process group does not, of course, reveal how the sampling distributions might change shape.

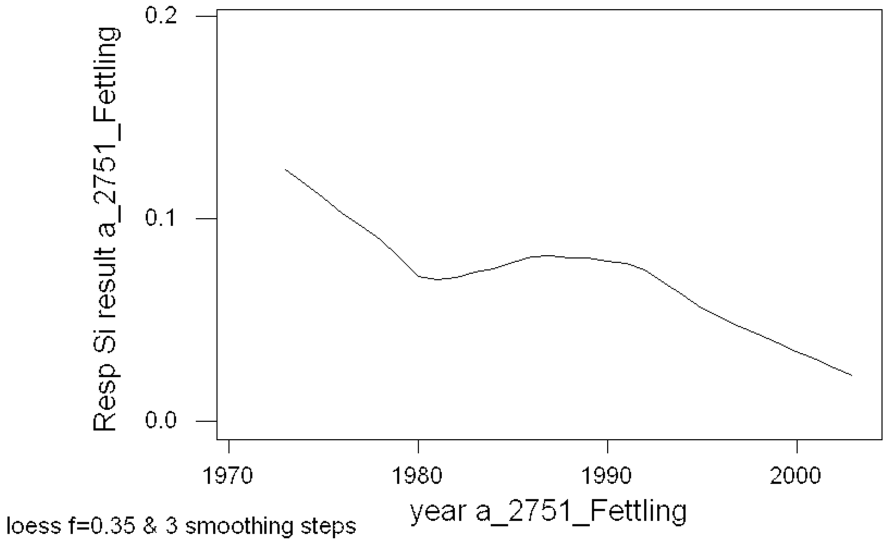


Figure 5.2 Lowess plot of the complete annual data sets for personal exposures to respirable silica of foundry workers involved in fettling processes, in the SIC category 2751, collected by the Alvechurch group of CTi.

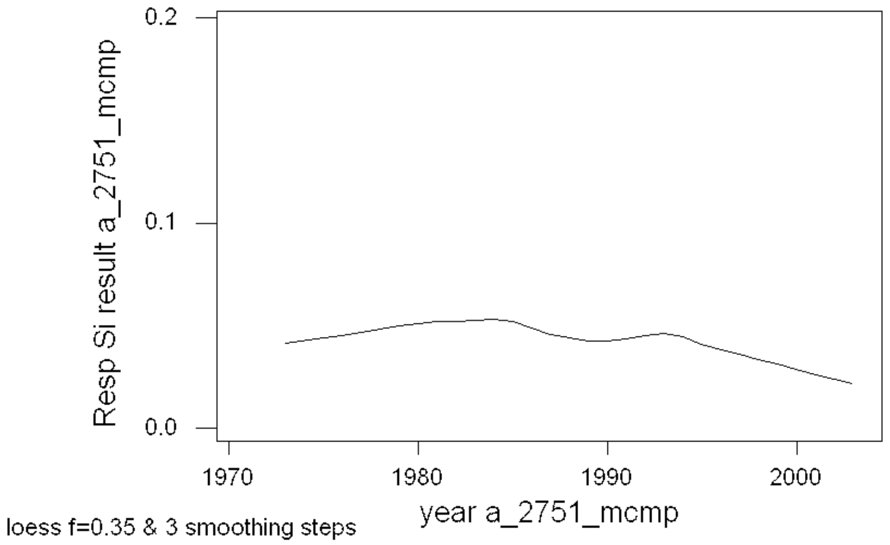


Figure 5.3 Lowess plots for the complete annual data sets for personal exposures to respirable silica for foundry workers involved in melting, moulding and core and mould production, in SIC category 2751, collected by the Alvechurch group of CTi.

Initially, I used a similar approach to investigate whether respirable silica exposure patterns differed over time for foundries of different sizes, categorised by the number of employees. Figures 5.4 to 5.7 illustrate this, demonstrating different patterns of time variation.

Of course, these plots involve using the complete annual exposure data sets and so are not able to reveal any possible changes in the shape of the sampling distributions with time. The patterns presented are difficult to interpret and, indeed, serve rather to confuse.

I considered that it was more useful to examine the lowess plots of the time variation of the upper fourths (similar to upper quartiles) for the two different groups of processes, examined earlier using the complete annual exposure data sets, namely: (a) fettling and (b) melting, core and mould production and mould production.. These are shown in Figures 5.8 and 5.9. These plots show more clearly the time variation that might be linkable with the impact of legislation and changes in regulatory limits.

The different patterns for the two process groups are likely to be explained by the difficulties in reducing exposures in the fettling process in the initial years following the introduction of the Health and Safety at Work, etc. Act. In this instance, the effect of foundry size is not considered but might be an important factor in terms of the implementation of better control. It could be that larger foundries might be better able to afford to improve dust control at an earlier stage than smaller foundries.

As we move further from the median towards the upper extremes, we might hope to see more pronounced effects that may differ in magnitude between the two process groups. So, I tested the idea further with lowess plots of upper eighths (E_U) and upper sixteenths (D_U). The upper eighths have 12.5% of the exposures above that threshold and the upper sixteenths have only 6.25% of the exposures above them. These are shown in Figures 5.10 to 5.13.

These seem encouraging results in that with both types of process groups investigated there have been significant reductions in the uppermost personal exposures to respirable dusts. Hence, we might expect there to be significant reductions in respirable silica, as well, although not necessarily in a proportionate manner.

An important message to emerge is that there is considerable value in undertaking routine monitoring of personal occupational exposures to hazardous substances and additionally that such monitoring can demonstrate the beneficial impact of legislation and tighter regulatory control in reducing occupational exposure risks.

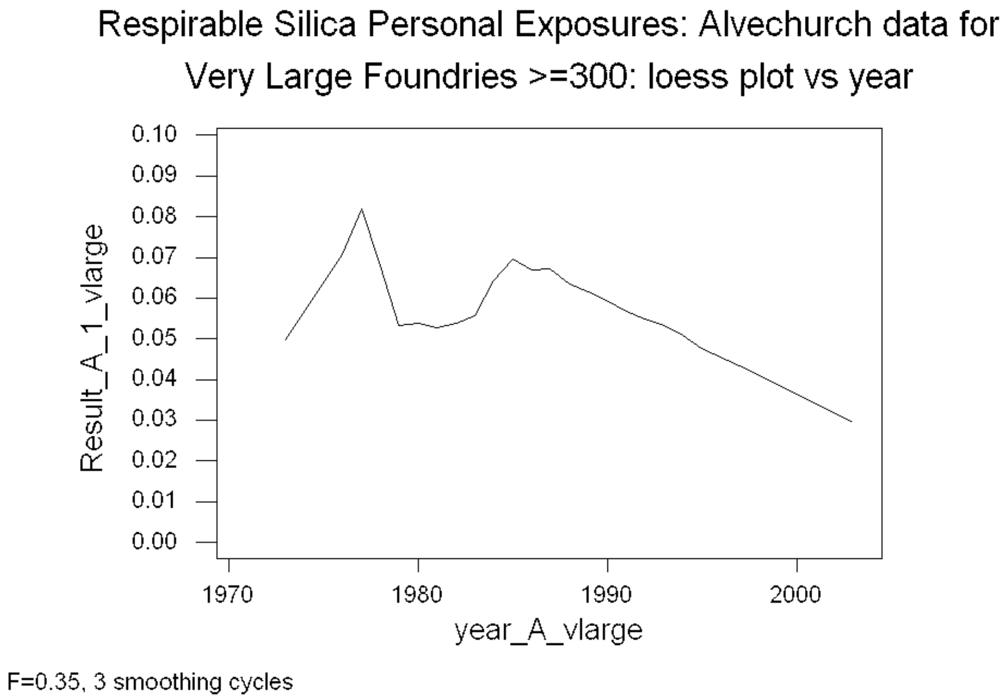


Figure 5.4 Lowess plot of personal exposures to respirable silica in very large foundries (>=300 employees), collected by the Alvechurch group of CTi.

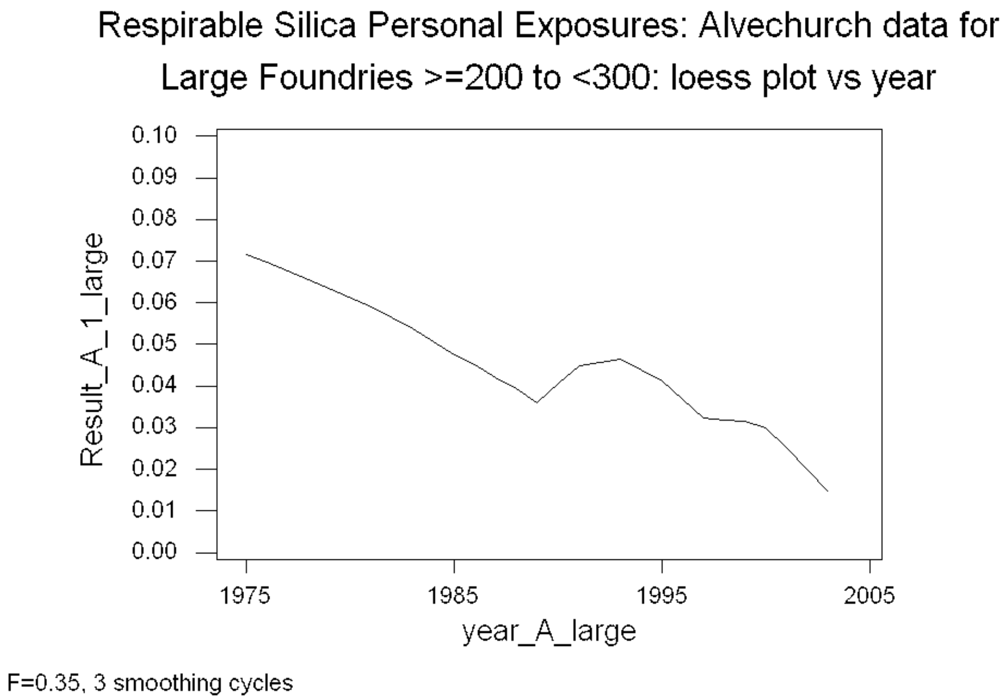
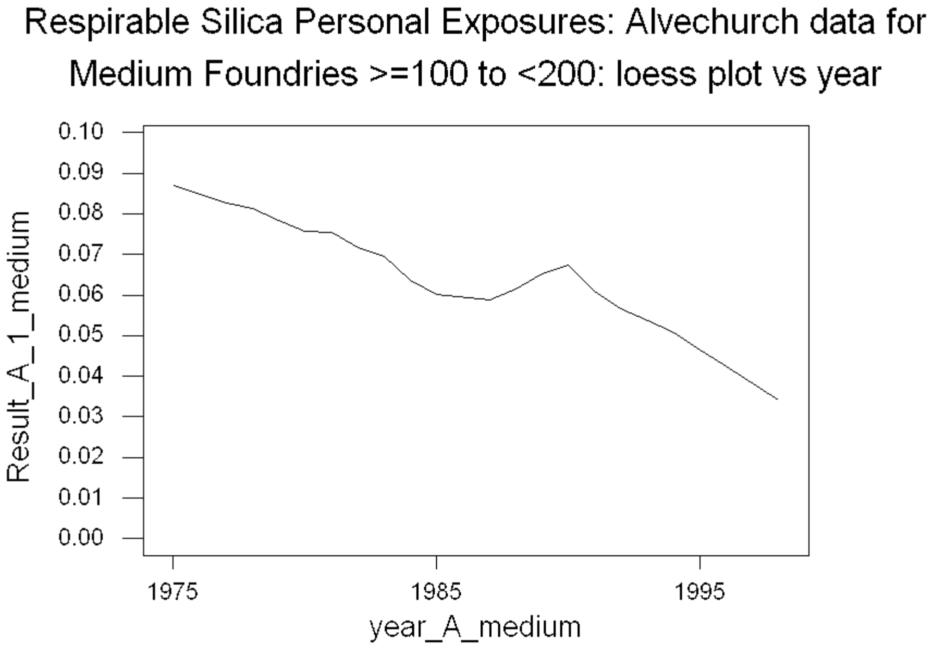
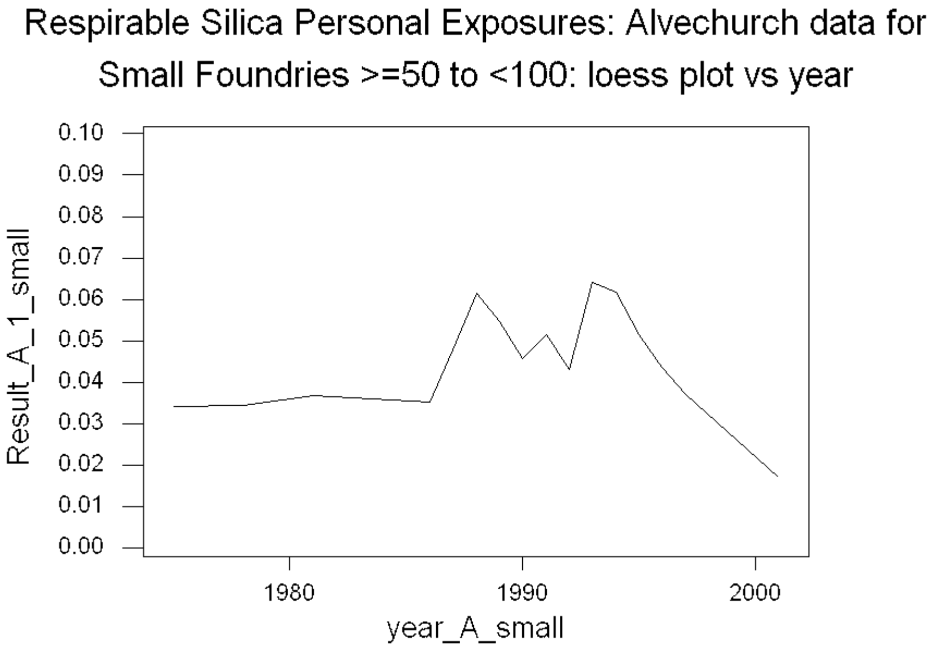


Figure 5.5 Lowess plot of personal exposures to respirable silica in large foundries (>=200 to <300 employees), collected by the Alvechurch group of CTi



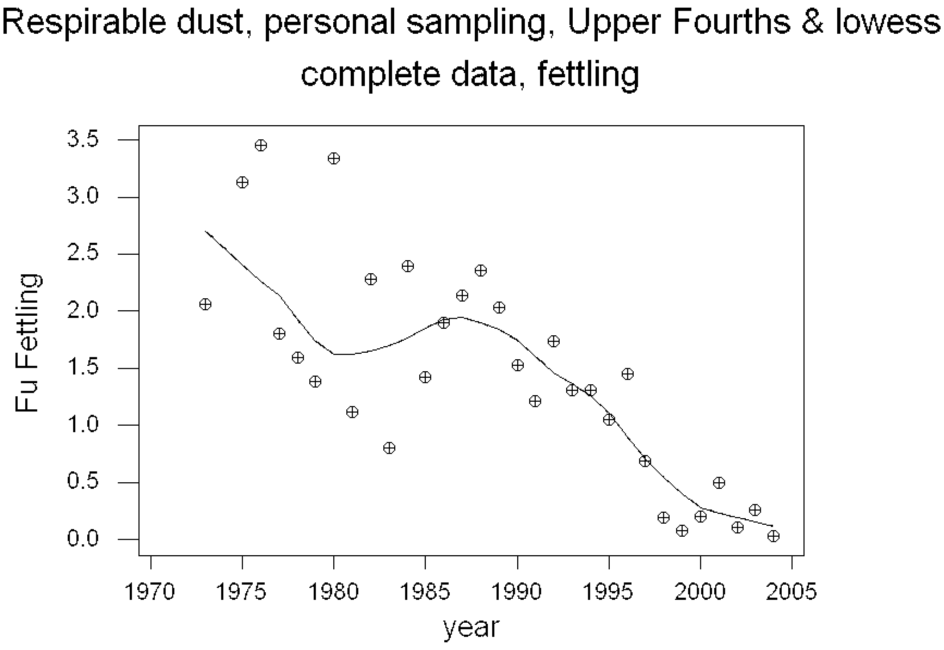
F=0.35, 3 smoothing cycles

Figure 5.6 Lowess plot of personal exposures to respirable silica in medium-sized foundries (>=100 to <200 employees), collected by the Alvechurch group of CTi.



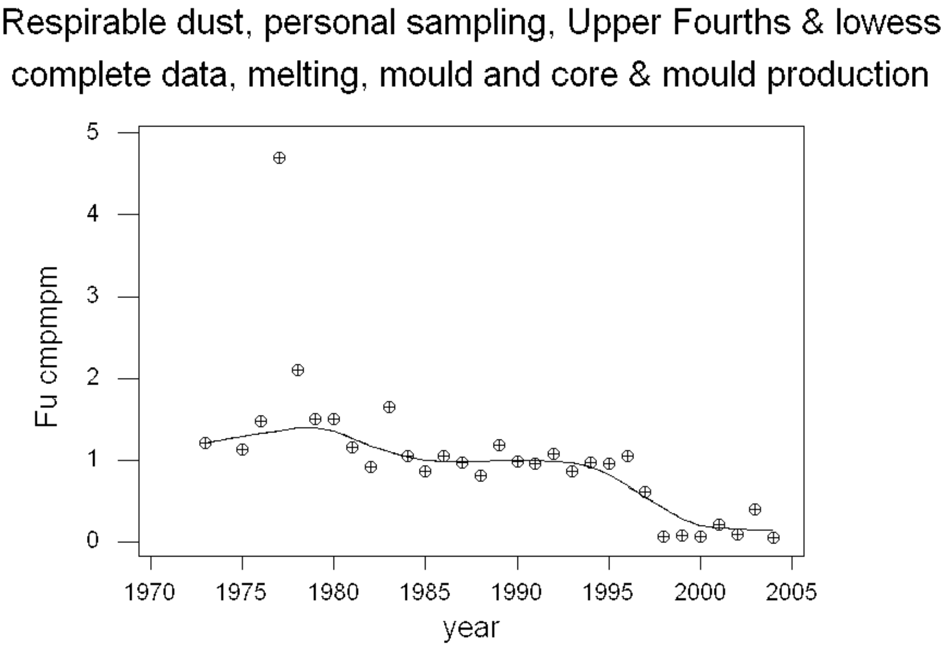
F=0.35, 3 smoothing cycles

Figure 5.7 Lowess plots of personal exposures to respirable silica in small foundries (>=50 to <100 employees), collected by the Alvechurch group of CTi.



F=0.35, 3 cycles

Figure 5.8 Lowess plots of the annual Upper Fourths of personal exposures to respirable dust by foundry employees involved in fettling, collected by both the Alvechurch and Sheffield groups of CTi.



F=0.35, 3 cycles

Figure 5.9 Lowess plot of the annual Upper Fourths of personal exposures to respirable dust of foundry employees involved in melting, mould and core and mould production, collected by both the Alvechurch and Sheffield groups of CTi.

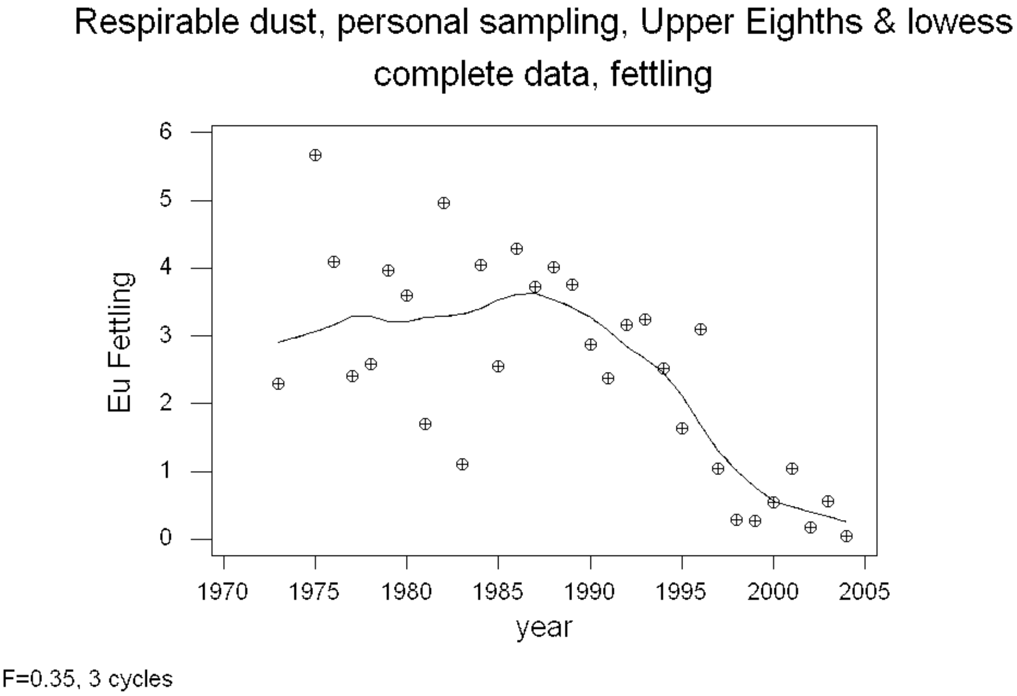


Figure 5.10 Lowess plots of the annual Upper Eighths of the personal exposures to respirable dusts by foundry employees involved in fettling processes, collected by both the Alvechurch and Sheffield groups of CTi.

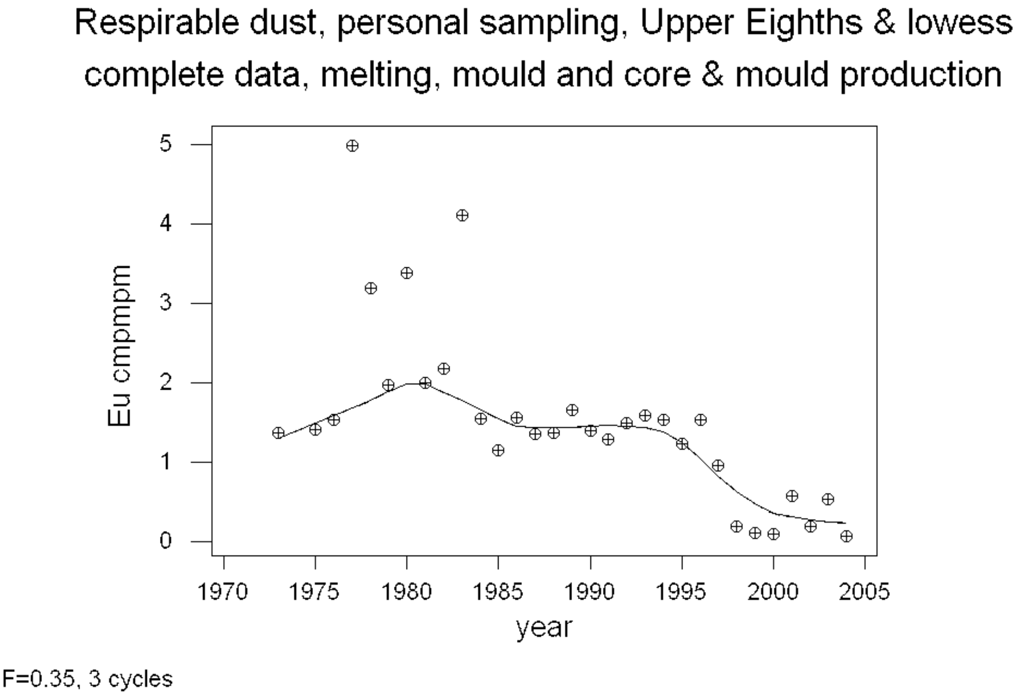


Figure 5.11 Lowess plots of the annual Upper Eighths of personal exposures to respirable dusts by foundry employees involved in melting, mould and core and mould production, collected by both the Alvechurch and Sheffield groups of CTi. Note that the Upper Eighths are often more extreme prior to 1984.

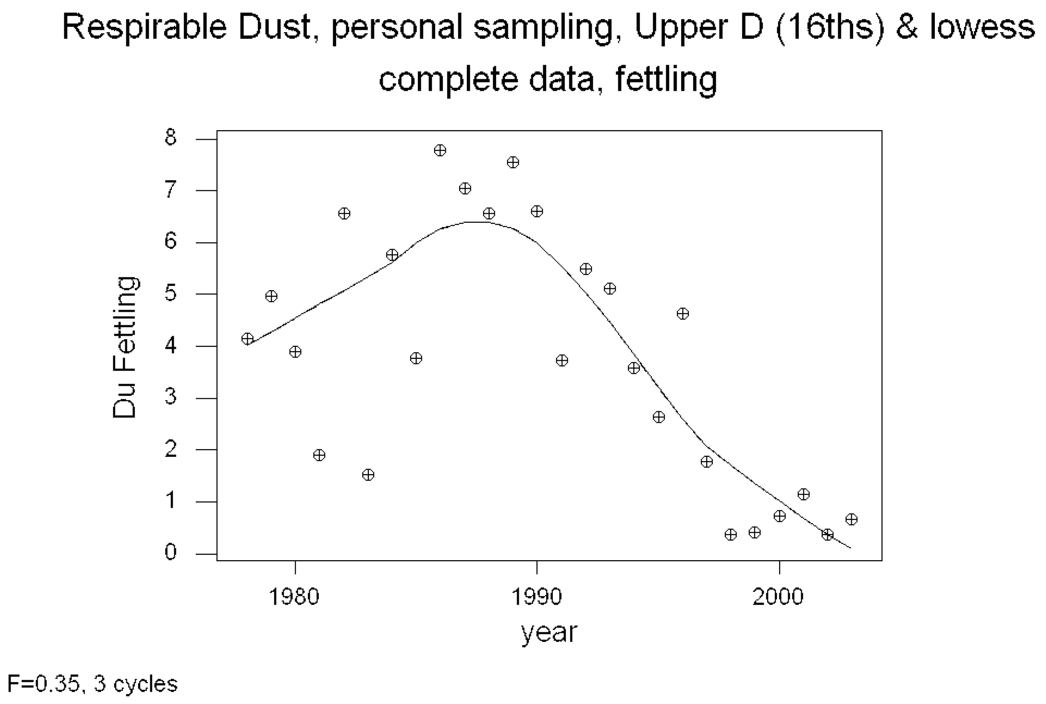


Figure 5.12 Lowess plot of the annual Upper Sixteenths of personal exposures to respirable dust of foundry employees involved in fettling processes, collected by both the Alvechurch and Sheffield groups of CTi.

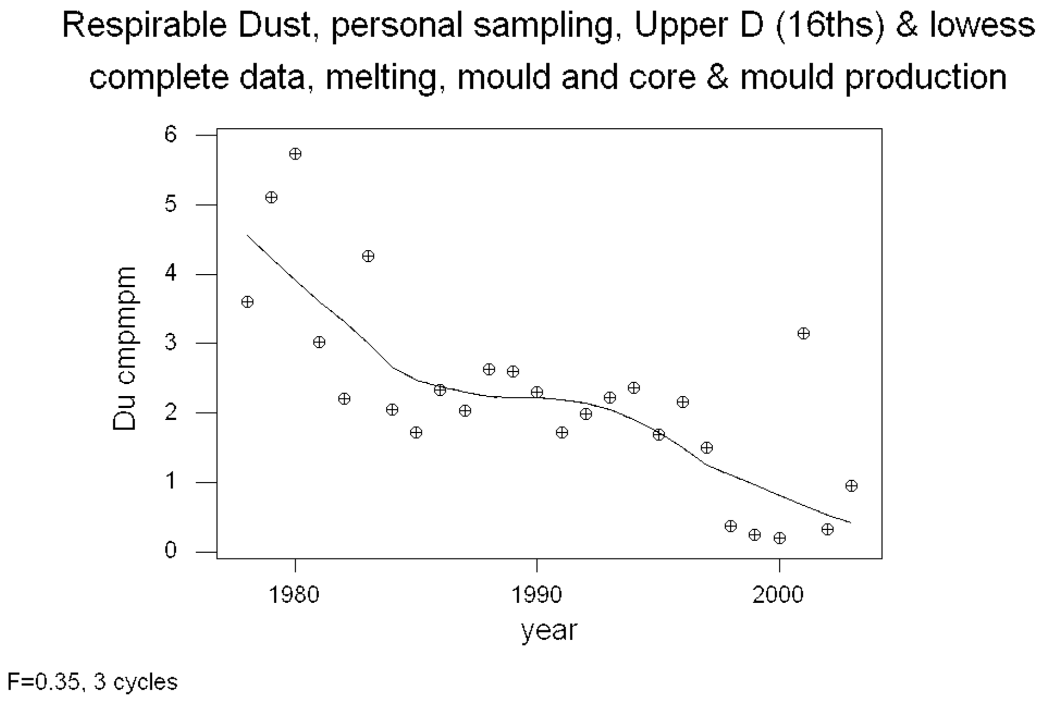


Figure 5.13 Lowess plots of the Upper Sixteenths of personal exposures to respirable dusts by foundry employees involved in melting, mould and core and mould production, collected by both the Alvechurch and Sheffield groups of CTi.

From the viewpoint of the impact of the Health and Safety legislation and tightening of regulatory exposure limits, the most interesting and significant observations from these plots, are the dramatic reductions in the upper parts of the distributions of the sampled exposures. Upper Fourths (roughly the same as Upper Quartiles), the Upper Eighths (approximately the upper 12.5 percentiles) and the Upper Sixteenths (approximately the upper 6.25 percentiles) for the very dusty fettling processes of respectively 89%, 92% and 94% over the period from 1987 to the present day. Even for the much less dusty processes in the mmpcmp category, there were respectively 82%, 79% and 82% reductions in exposures. Over the period since 1985, I found similar magnitudes of reductions. For example, I noted that there was around a 70% reduction in the mean level of the Upper Sixteenths (approximately the 93.5 percentile) of the personal annual exposures in the personal exposures to respirable silica, as judged by the Cusum Span test (see the last Figure in Chapter 4).

What is very disappointing to observe is the considerable reduction, in recent years, in the amount of personal exposure monitoring by the foundry industry itself. In terms of good governance in respect to health and safety of employees, this raises serious concerns that proper judgement cannot be made about the sustainability of the excellent improvements in exposure risks achieved by the foundry industry in previous years.

Thus the role of the HSE as the regulator of occupational risk and an important guardian of public health should ideally come into play in fulfilment of its obligations under its Public Service Agreement, as I remarked in Chapter One and as I feel obligated to reinforce and reiterate here.

Bibliography and References

Bowman, A. W. and Azzalini, A.: “Applied Smoothing Techniques for Data Analysis” 1997, Oxford University Press, Oxford, UK

Box, G. and Luceño, A.: “Statistical Control by Monitoring and Feedback Adjustment” 1997, John Wiley and Sons, Inc., New York, N.Y., U.S.A.

BS5703: “Guide to data analysis and quality control using CuSum techniques” in four parts. (1980 - 1982) British Standards Institution, London, U.K.

BS5703:2003: “Guide to data analysis and quality control using cusum techniques” in four parts (2003) British Standards Institution, London, U.K.

Chen, Z., Bai, Z. and Sinha, B. K.: “Ranked Set Sampling – Theory and Applications” 2004, Springer-Verlag New York Inc., New York, N.Y., U.S.A.

Cleveland, W. S.: “The Elements of Graphing Data” 1994, Hobart Press, Summit, New Jersey, U.S.A.

Efron, B. and Tibshirani, R. J.: “An Introduction to the Bootstrap” 1993, Chapman and Hall, London, U.K.

Gibbons, J. D.: “Nonparametric Methods for Quantitative Analysis” 1976, Holt, Rinehart and Winston,New York, N.Y., U.S.A.

Harvey, R. P. in Cullis, C. F. and Firth, J. G., eds.: “detection and measurement of hazardous gases” 1981, Heinemann Educational Books Ltd., London, U.K., Chapter 6 “Statistical aspects and air sampling strategies”

Hoaglin, D. C., Mosteller, F. and Tukey, J. W., eds.: “Understanding Robust and Exploratory Data Analysis” 1983, John Wiley and Sons, Inc., New York, N.Y., U.S.A.

Hoaglin, D. C., Mosteller, F. and Tukey, J. W., eds.: “Exploring Data Tables, Trends and Shapes” 1985, John Wiley and Sons, Inc., New York, N.Y., U.S.A.

Kristensen, F. B., Hørder, M. and Poulsen, P. B., eds.: “Health Technology Assessment Handbook” 2001, Danish Institute for Health Technology Assessment (www.dihta.dk)

Leidel, N., Busch, K. A. and Lynch, J. R.: “Occupational Exposure Sampling Strategy Manual” 1977, DHEW (NIOSH) Publication No. 77-173, Superintendent of Documents, U.S. Government Printing Office, Washington D.C., U.S.A.

Mosteller, F. and Tukey, J. W.: “Data Analysis and Regression – a second course in statistics”1977, Addison-Wesley Publishing Company, Inc., Reading, Ma., U.S.A.

NHS HTA Annual Reports, available from the NHS website

Sprent, P.: “Data Driven Statistical Methods” 1998, Chapman and Hall, London, U.K.

Sprent, P. and Smeeton, N. C.: Applied Nonparametric Statistical Methods” 2001, 3rd. ed., Chapman and Hall, London, U.K.

Taylor, A. L., Tait, S. P., Porter, M. A., Perry, M. J. and Nicolson, R. W.: "Automatic breakpoint detection for retrospective cumulative sum charts" Pharmaceutical Statistics, 2002, **1**: 25-34.

Thompson, J. M. in Tranter, R., ed.: "Design and Analysis in Chemical Research" 2000, Sheffield Academic Press, Sheffield, England. Chapter 5 "Robust resistant and nonparametric methods"

Thompson, J. M., McElvenny, D. and Llewellyn, D.: Workshop on "Statistical methods in regulatory occupational hygiene risk assessment" British Occupational Hygiene Society Annual Scientific Meeting, Imperial College, London, April 10th 2003.

Thompson, S. B. and Seber, G. A.: "Adaptive Sampling" 1996, John Wiley and Sons, Inc., New York, N.Y., U.S.A.

Wells, D. and Greenall, A.: personal communications, 2004.

Williamson, R. J.: "Span tests on Cusums" The Statistician, 1985; **34**: 345-356.

Woodward, R. H. and Goldsmith, P. L.: "Cumulative Sum Techniques" Mathematical and Statistical Techniques for Industry Monograph No. 3, 1964, Oliver and Boyd Ltd., Edinburgh.

Printed and published by the Health and Safety Executive
C1.10   07/05

HSE
BOOKS

MAIL ORDER
HSE priced and free
publications are
available from:
HSE Books
PO Box 1999
Sudbury
Suffolk CO10 2WA
Tel: 01787 881165
Fax: 01787 313995
Website: www.hsebooks.co.uk

RETAIL
HSE priced publications
are available from booksellers

HEALTH AND SAFETY INFORMATION
HSE Infoline
Tel: 0845 345 0055
Fax: 0845 408 9566
e-mail: hseinformationservices@natbrit.com
or write to:
HSE Information Services
Caerphilly Business Park
Caerphilly CF83 3GG

HSE website: www.hse.gov.uk

RR 373

£10.00

ISBN 0-7176-6150-4

9 780717 661503