OCCUPATION-RELATED SUICIDE

The impact of the Census Act and the Population (Statistics) Act on research based on individual-level mortality data.

LAUREL M MILWID

Institute for Social, Cultural and Policy Research
University of Salford,
Salford, U.K.

Submitted in partial fulfilment of the requirements of the degree of
Master's in Research in Social Sciences (Sociology)

June 2004

milwid@googlemail.com

Published in 2011 through Lulu Press, Inc.
www.lulu.com

ISBN: 978-1-4709-0630-6

CONTENTS

LIST OF TABLES

ACKNOWLEDGEMENTS

The advice, support, and guidance of my research supervisor Greg Smith, were of immeasurable value during the course of this project. I would also like to acknowledge the advice, support and encouragement that Steve Barrow offered so readily throughout, as well as that of Neal Hazel.

Staff at the Centre for Longitudinal Study User and Information Support (CeLSIUS) and the Office for National Statistics (ONS) were generous with support and information, which I acknowledge gratefully. In particular, I would like to acknowledge the generosity of advice and information so readily given by Folkert Van Galen of ONS.

My children, Lara and Matthew, were equally generous in their support, and their encouragement was of immeasurable value.

ABSTRACT

A comprehensive review of the literature reveals a deficit in explanation for the differences in suicide-risk between occupations. Two studies of occupation-related suicide in England and Wales (Kelly et al, 1995; Kelly & Bunting, 1998) are unusual in reporting suicide-data for multiple occupations. These studies rank a series of occupations according to their associated suicide risk, which is measured in terms of the Provisional Mortality Ratio for suicide. The studies do not incorporate controls for demographic covariates of occupation, apart from gender. An American study (Stack, 2001), by contrast, is based on a logistic regression model of analysis, with controls introduced for age, gender, marital status, and race. Comparing the bivariate logistic regression odds ratios with the respective multivariate logistic odds ratios for a series of occupations demonstrates the importance of incorporating demographic controls in the assessment of occupational risk of suicide.

Examination of the overlap between the findings of these studies demonstrates a series of inconsistencies, and suggests that this may partly due to different methodologies. This highlights the importance of rendering studies to a common theoretical base in order to facilitate valid comparisons, and provides justification for development of a statistical model analogous to that employed in Stack's American study, with the objective of testing it using mortality data for England and Wales. Testing this model depends for its viability on access to individual-level (as opposed to aggregate) mortality data. The only potential source is identified as the Longitudinal Study, which links Census and vital-event data for 1% of the population of England and Wales from 1971, and is held by the Office for National Statistics (ONS). The Centre for Longitudinal Studies Information and User Support (CeLSIUS) serves as the user gateway to the Longitudinal Study. My application for access to the Longitudinal Study database revealed that Census-based data and vital events data (birth and death registration, widow(er)hoods, cancer registrations, migration, enlistments, entries to long-stay hospitals) are protected in England and Wales by the Census Act and the Population (Statistics) Act. Secondary analysis of this data is restricted to officers of the Office of National Statistics. In the United States, by contrast, access to population data (such as the National Mortality Detail Files, which served as the source of data for Stack's study) is not only permissible, but the accessibility of such data (which is held by the National Centre for Health Statistics) to the public is *mandated*.

A discussion of the respective legislation underlying the different policies in England and Wales and the United States relating to privacy of mortality data highlights a trade-off between the privacy of the dead and social scientific research, and gives rise to the question: To what extent do the dead *need* protection?

Note on referencing of websites and web-pages

Websites and webpages are referenced throughout this thesis in the following format, and are listed separately in the Reference section:

www.CDC[2]	Centre for Diseases Control and Prevention (CDC). *Self inflicted injury / suicide.* www.cdc.gov/nchs/fastats/suicide.htm

CHAPTER 1

INTRODUCTION

1. BACKGROUND

In my work in Zimbabwe in the late 1990s as assistant director of a human rights agency and as a counsellor, I became increasingly aware of the influence that macrosocial factors have on individual psychological functioning. The traditional psychotherapeutic models that were favoured in Zimbabwe at the time did not, in my view, adequately recognise the influence on individual psychological health of the economic, social and political factors that prevailed, such as spiralling inflation, redistribution of commercial farms, high rates of unemployment, and the precarious state of the rule of law. Even though clients might meet the diagnostic criteria for diagnosis of depression or anxiety, clinical assessment would frequently reveal that structural factors, such as inflation resulting in dramatic erosion of income, were more significant influences on the clients' psychological status than were those factors traditionally assumed to be determinants for depression or anxiety, such as deprivation of affection or dysfunctional relationships.

I increasingly felt that it was unrealistic to attempt to understand an individual's psychological status without reference to the prevailing political, economic, and human rights climate, particularly in a country that is in a state of economic and judicial disarray. I believed that a more realistic interpretation of their 'condition' depended on recognising the influence of 'social' pathology (i.e. dysfunction in the social and political structure of society) rather than mechanistically ascribing their psychological distress to 'psycho' pathology.

In trying to formulate the research question for this thesis, my task was to attempt to translate the perceptions described in the previous paragraph into a local context. While it is true that the United Kingdom hosts a relatively stable economy and is committed to democratic political process, I do not believe that 'democracy' necessarily means that political and social institutions are innocent of exerting a malevolent influence on the psychological functioning of individual citizens. In Zimbabwe and other developing countries, individual rights are blatantly and transparently violated. But the notion of 'freedom' in the United Kingdom is, in my view, a semantic bluff, and the dignity and psychological status of citizens of the United Kingdom are almost as vulnerable to violation as are the dignity and

psychological status of citizens of Zimbabwe. The primary difference is that, in the United Kingdom, the process of violation is disguised. Social and political structures colonise individual lives and psyches insidiously, making the state, together with its agents, a 'manipulator' of significant proportions. Policy concerning educational 'inclusion' for example is, on the face of it, benevolent. But, by assigning the disruptive student the fundamental right to be 'included' in the classroom, it imposes at the same time the less democratic obligation on the teacher to 'include' him (or her), together with the obligation to accommodate his (or her) abuse. In the case of the teacher presenting with 'depression', a psychotherapeutic approach that fails to acknowledge the impact of institutional factors on his (or her) psychological status is inadequate. The fundamental 'problem' arises in social policy, not in his (or her) psyche. His (or her) psychological status is merely a symptom of 'pathology' in the social structure.

2. THE RESEARCH HYPOTHESIS

The introductory paragraph of this chapter describes the background to my interest in the influence of conditions in the wider social environment, on the psychological well-being of the individual. My primary hypothesis is that macro-social structures, particularly those that are not autonomous from party political manipulation, manifest a malignance that exerts a significant and measurable negative influence on the intrapsychic functioning of individuals whose personal or professional lives are subject to the authority or influence of those systems. Psychological assessment of these individuals is likely to reveal that disturbances of psychological wellbeing are more (or, at least, as much) a manifestation of 'social' pathology in the macrosystem as of 'psycho' pathology in the individual.

In seeking to operationalise this research hypothesis, it seemed more practical to narrow the field of study down to investigation of relationships within a single macrosocial structure. The fact that I was, at the time, teaching in a high school influenced me to focus my study on the education system and to structure the research question in terms of investigating the relationship between the 'education system' and the psychological status of 'teachers', as one category of individuals whose lives are influenced by the education system. If a relationship could be demonstrated, I would then consider the question of whether this relationship could be generalized to other macrosocial structures.

In the first stages of drafting the research proposal, I explored the idea of testing this relationship by using psychometric testing to evaluate negative affect in teachers, and then using questionnaires to evaluate their perceptions of which factors they considered to be causally related to their negative affect. However, on review, I realised that there would be numerous practical, ethical, professional and

economic barriers to conducting psychological testing for the purposes of evaluating negative affect, and a definitive indicator of negative affect would be more appropriate. This was the cue for orienting the research question towards an analysis of suicide rates. 'Suicide' is an unequivocal indicator of negative affect, and provides a series of advantages in relation to measurement of affect in subjects. These include eliminating the need for interviewing subjects, eliminating the need to develop an instrument for measuring the psychological distress associated with teaching, and eliminating the need to derive a representative sample of teachers.

I initially proposed to investigate patterns of 'suicide' (as an unequivocal indicator of psychological distress) in the teaching population, and to explore the hypothesis that suicide, as an extreme indicator of compromised psychological wellbeing, is related to manipulation of the education system by the State, in the interests of promoting political power and at the expense of consideration for the dignity and well-being teachers.

At about this time, the media reported the suicide of David Kelly, who had been senior advisor on biological warfare for the United Nations in Iraq until 1999 and who was, at the time of his death, the scientific advisor to the Proliferation and Arms Control Secretariat of the Ministry of Defence (www.BBC). Kelly was alleged to be the main source for allegations by the BBC that Downing Street had 'sexed up' the September 2002 dossier on Iraq's weapons of mass destruction, which Downing Street had used to support its attempt to justify England and Wales's invasion of Iraq in March 2003.

After his name was leaked to the press, Kelly found himself at the centre of a public row between the BBC and Downing Street. As the central figure in the enquiry conducted by the Foreign Affairs Committee, he did not have an avenue for manouevre and, although a man who normally coped well under pressure (www.BBC), sought recourse to the ultimate means of escape: his own death.

This was a case where the connection between 'suicide' and 'government manipulation' was blatantly apparent, and it prompted me to consider widening the scope of my project beyond the education system and to explore the relationship between suicide and occupations in general.

The theoretical literature does not consistently support a positive correlation between occupation and the risk of suicide, and reveals varying and often contradictory explanations of relationships between suicide, occupation, and occupational conditions (such as social status) that may influence suicide patterns (e.g. Powell, 1958; Lampert et al, 1984), access to means for committing suicide (e.g Kelly & Bunting, 1998; Hawton et al, 1998), client dependency (e.g. Labovitz & Hagedorn), social integration (e.g. Hood-Williams, 1996), and seasonal cycles (e.g. Wenz,1977). The aim of this research project was to explore this literature, with the goal of developing a research question which would form the basis for an empirical study. Informed by an often contradictory literature on the relationship

between occupation and suicide, the objectives of this project are itemised as follows:

i) To evaluate studies of the relationship between occupation and the risk of suicide with the aim of identifying gaps in the theoretical literature that warrant further investigation;
ii) To develop a research question for investigation;
iii) To conduct an empirical study designed to address the identified question.

These objectives will be addressed with reference to the following concepts:

i) 'Suicide' as an indicator of psychological distress
ii) All members of 'occupations in general' as the research population.

And, finally, central to this thesis is the primary hypothesis (as outlined earlier in this section) that:

"*Macro-social structures, particularly those that are not autonomous from party political manipulation, manifest a malignance that exerts a significant and measurable negative influence on the intrapsychic functioning of indivuduals whose personal or professional lives are subject to the authority or influence of those systems.*"

The process whereby this hypothesis was addressed, the research questions that emerged, the methodology around which this thesis is centred, and the outcomes of the project, are overviewed in the following section.

3. THESIS STRUCTURE

This thesis consists of 6 chapters. Following this introductory chapter are two chapters (2 and 3) out of which the theoretical concerns which are the focus of this thesis will emerge. The fourth and fifth chapters centre round methodological issues and the nature and availability of primary data for analysis, while the sixth chapter offers a retrospective evaluation of the project.

3.1 CHAPTER 1

Chapter 1 has outlined the background to this thesis. The hypothesis that forms the basis for this thesis has been identified, as has the conceptual basis for development of a research question. The aims of the thesis have been presented, and these aims inform the overall structure of the thesis.

3.2 CHAPTER 2

Chapter 2 provides an overview of theoretical literature relevant to analysis of the relationship between occupation and suicide. Two key studies (Kelly et al, 1995; Kelly & Bunting, 1998) are identified which are unusual in England and Wales by virtue of being based on an analysis of *multiple* occupations. An American study (Stack, 2001) is identified, which is also based on an analysis of *multiple* occupations, but which contrasts with studies in England and Wales in relation to underlying theoretical assumptions, methodology, and the *measure* of suicide risk that is used.

The American study is based on a logistic regression model of analysis, with controls introduced for four demographic covariates of occupation (age, gender, race, marital status). Suicide risk is measured, on the basis of this model, in terms of logistic regression odds ratios. Studies in England and Wales, by contrast, do not introduce demographic controls (apart from gender) into the analysis, and the measure of suicide risk is the Provisional Mortality Ratio.

The emphasis by Stack (2001) on the importance of controlling for demographic factors, the evidence that a similar study has not yet been conducted in England and Wales, and the evidence that only very few studies on *multiple* occupations have been conducted in England and Wales, suggest a 'gap' in the literature and served to provide the focus for development of a research question for my own study.

3.3 CHAPTER 3

This chapter provides discussion of the methodology on which the two England and Wales studies identified in Chapter 1 is based, and the theoretical assumptions which underlie this methodology. A similar discussion is offered of Stack's study.

Discussion in this chapter highlights the importance of controlling for the demographic covariates of occupation, a consideration which is central to the theorising outlined by Stack. This section also examines the overlap between the findings of Kelly & Bunting's study and those of Stack's, and demonstrates the inconsistencies that arise as a result of the use of different methodologies. This discussion serves to highlight the complexities involved in assessing the reliability and validity of findings from investigations of occupation-related suicide by comparison with results from other studies, and demonstrates the importance of rendering studies to a common theoretical base before valid comparisons can be made.

The observations detailed in the previous chapter are put forward as justification for development of a research proposal which has as its objective an

attempt to replicate the methodology used by Stack, using local mortality data. An attempt to identify an appropriate source of primary data for this analysis serves as the focus for the next chapter.

3.4 CHAPTER 4

An investigation of occupation-related suicide risk that is based on a multivariate model of analysis analogous to that employed by Stack depends, fundamentally, on access to individual-level mortality data (as opposed to aggregate data). This data must include information about gender, marital status, age and race. In Chapter 4, potential sources of such data are explored. A link between agencies that hold mortality data for England and Wales is identified, and the apparent supremacy of the Office for National Statistics (ONS) in controlling this data is demonstrated. The legislative basis for this supremacy is identified (and addressed more fully in Chapter 6).

The Longitudinal Study (which is held by ONS) is identified as the only potential source of individual-level mortality data for England and Wales, and the Centre for Longitudinal Study Information and User Support (CeLSIUS) is identified as the gateway for researchers who wish to apply for access to data held on the Longitudinal Study.

In terms of the legislative restrictions referred to in an earlier paragraph, external users may not access individual-level data held on the Longitudinal Study. However, a conversation between one of my academic advisors and a member of CeLSIUS staff led to the impression that the conflict between my requirement to base the analysis on individual-level data and ONS's policy not to permit direct access to this data could potentially be resolved by CeLSIUS facilitating *indirect* access to the data by conducting the statistical analyses on their own computers in terms of specifications that I would provide. It was on the basis of this possibility that I submitted a proposal to CeLSIUS for access to mortality data held on the Longitudinal Study.

3.5 CHAPTER 5

The response from CeLSIUS centred on suggestions for modifying the research proposal in such a way as to allow the analysis to be based on *aggregate* data. Although the response did not address the request for indirect access to individual-level data directly, the impression was given that this request was not being entertained. Since the viability of a project designed round a logistic regression model of analysis depends fundamentally on access to individual-level data, the goal of attempting to replicate Stack's multivariate model of analysis of occupation-related suicide risk ultimately proved to be unattainable.

As Chapter 5 demonstrates, the Longitudinal Study appears to be the primary venue of mortality data for England and Wales for researchers but, at the same time, a futile venue for *individual*-level mortality data, and (in the case of studies of occupation-related suicide) a fundamentally limited venue for *aggregate* mortality data.

3.6 CHAPTER 6

Chapter 6 provides an overview of the legislative complexities that underlie protection of census-based and vital-events (including mortality) data in England and Wales. It also examines differences in legislation that make statistical analysis of census-based data and vital-events data possible in the United States but unachievable (for the public) in England and Wales. It is these differences that enable studies such as Stack's to be conducted in the United States, but make them fundamentally unfeasible for England and Wales, unless they are both initiated from within and conducted within the Office for National Statistics. As this chapter suggests, this fundamental curtailment of mortality-related research has clear implications for research and for the potential of research to inform social policy.

4. CONCLUSION

Restrictions on access to relevant mortality data proved to be the deciding factor on the viability of the research proposal outlined in this thesis. But, even though the proposal, as it stood, ultimately proved to be unfeasible, this thesis contributes to the existing theoretical literature on the relationship between occupation and suicide in a number of ways:

i. This study provides a review of the theoretical foundations of and empirical evidence for hypotheses relating occupation to suicide;
ii. It proposes an empirical model to test risk for suicide on the basis of occupation;
iii. It identifies a series of topics related to suicide and occupation that could justifiably provide the basis for future research initiatives;
iv. It addresses the issue of 'privacy' and outlines the restrictions that legislation imposes on access to mortality data;
v. It identifies a question that is fundamental to discussion of the implications of the Census Act and the Population (Statistics) Act for research and public policy relating to mortality and occupation: *To what extent do the dead need protection?*

CHAPTER 2

AN OVERVIEW OF THE LITERATURE

1. INTRODUCTION

Suicide is, within the context of Western society, a powerful indicator of negative affect, and is a subject that has been of enduring interest to sociologists. The founder of the scientific study of suicide is generally regarded to be Emile Durkheim, whose seminal work (Durkheim, 1952) emphasises the influence of macrostructures on micro-level phenomena, a phenomenon to which I referred in Chapter 1, section 2. Durkheim's monograph, which was first published in 1897, continues to the present day to be an important reference for suicide studies (Peloso, 1992). The relevance of Durkheim's work to the theme of this thesis, i.e. occupation-related suicide, is discussed in this chapter. This discussion is followed by an overview of secondary sources which have specific relevance to development of the research question that will be the focus of this thesis, and which will start emerging towards the end of this chapter.

As the second section of this chapter will demonstrate, a review of secondary sources reveals that a range of factors operating at both individual and demographic level serve to either predict or protect against suicide. Occupation is one such predictor factor. But the relationship between occupation and suicide is complex, because of the influence of the multitude of other predictor and protector factors, as well as the influence of longitudinal trends in suicide in the general population.

An American study (Stack, 2001), which is a central focus of this thesis, posits a *multicausal heuristic model of occupation-related suicide* and conceptualises the relationship between occupation and suicide in terms of four major factors which contribute to the suicide risk associated with any particular occupation. This model will be discussed in Section 3 of the present chapter, with a particular focus on the demographic dimension of the model. A multivariate analysis of the demographic dimension of this model forms the basis of the methodology for Stack's study (Stack, 2001), which is central to this thesis (see Chapter 3, section 2) and which provides the basis for construction of my own research question.

Section 4 of this chapter offers a discussion of the influence on occupation-related suicide trends of suicide trends in the general population. As this section will demonstrate, with reference to local studies, longitudinal trends in suicide (such as demographic transformations of both the general population and of individual

occupations) tend to enhance the complexity of any analysis of occupation related suicide, which means that evaluation of occupational trends in suicide should be examined against the background of trends in suicide in the general population.

Section 5 offers a discussion of four factors that may influence the recording of primary suicide data and which, by influencing suicide statistics, may have a bearing on statistical analyses relating to occupation related suicide, namely: conventions for classifying suicide as a cause of death, mortality coding systems, the Suicide Act of 1961, and conventions for classifying occupations.

The chapter closes by briefly contrasting two key studies in England and Wales on occupation-related suicide (Kelly et al, 1995; Kelly & Bunting, 1998) with the American study (Stack, 2001) which has already been referred to in this introduction. This closing discussion points to an apparent gap in the theoretical literature on occupation-related suicide in England and Wales, and it is this gap that informs the formulation of a research question for this thesis.

2. AN OVERVIEW OF THE LITERATURE

Durkheim's classic study 'Suicide', which was initially published in 1897, was the first outstanding contribution towards explanation of variability in suicide rates in terms of social, cultural, and demographic variables (Gibbs & Martin, 1958). In this study, Durkheim asserted that certain social phenomena exert a powerful influence over individual behaviour (Hood-Williams, 1996), and that "the most impressive example is that of social forces which drive individuals to their deaths, each believing that he is obeying only himself" (Raymond Aron, quoted in Taylor, 1982: 21). He argues that the act of suicide is a manifestation not of a victim's "personal temperament" but of a social condition external to his or her private experience (Durkheim 1952: 299). The 'social condition' to which Durkheim refers is the existence of three suicidogenic forces (egoism, altruism, and anomy) which characterise the macrostructure of society. Each social group in society has, according to Durkheim (1952: 299), a "collective inclination for the act [of suicide]", an inclination which exists as a result of these suicidogenic forces. Durkheim describes these forces extensively in his monograph. But, in light of the focus in this thesis on occupational groups, it is the concept of 'collective inclination' towards the act of suicide that is of specific relevance.

Durkheim (1952) demonstrates his acknowledgement of the influence of occupation on suicide rates by comparing the rates of suicide between five occupational categories (trade, transportation, industry, agriculture and liberal professionals) for eight European countries, and thus demonstrating varying 'collective inclinations' towards the act of suicide. He concludes that industrial functions and commercial functions "furnish the greatest number of suicides", that

they are "almost on a level with the liberal professions", and that these three categories are "more afflicted than agriculture" (Durkheim, 1952: 257).

Also of particular relevance to the subject matter of this thesis is Durkheim's statistical conceptualisation of suicide risk. This conceptualisation centres on the "coefficient of aggravation" and the "coefficient of preservation" which Durkheim developed to serve as comparative population (or subpopulation) indicators for suicide risk (Durkheim, 1897: 177; www.TU). This statistic indicates how many times less frequent suicide is in one group that in another of the same age, and is a measure of the degree to which an independent variable (e.g. marriage) has a statistically negative (or 'protective') effect on the suicide rate, as shown in the following fictitious example, for men aged 30-39 in the UK:

$$\frac{\text{Suicide rate for unmarried men aged 30-39 in UK (145)}}{\text{Suicide rate for unmarried men aged 30-39 in UK (88)}} = \text{Marital coefficient of preservation for men aged 30-39 in UK (1.64)}$$

If the coefficient (in this case known as the 'marital' coefficient of preservation) is greater than 1, then that variable (in this case 'marriage') reduces the suicide rate for the relevant population group. The larger the coefficient, the greater the effect of the variable. When the coefficient of preservation falls to below 1, then the variable increases the suicide rate and is called a 'coefficient of aggravation' (Durkheim, 1897: 235; www.TU).

Durkheim applied this method of analysis to statistics for suicides in the military and civilian populations (Durkheim, 1897: 235). By computing the military coefficient of aggravation for soldiers in proportion to civilians for the age-group 20-30 for five countries, (France, Prussia, England, Italy and Austria) he demonstrated that the rate of suicide amongst soldiers is greater than for civilians of the same age-group, for each of the countries included in the analysis. His discussion of this topic serves to highlight the ongoing relevance of his methodology to contemporary studies of occupation-related suicide.

Of further relevance to contemporary studies and, in particular, to discussion of the methodology employed by Stack, which incorporates demographic controls in the analysis of occupation-related suicide (and which will be discussed more fully later in this chapter), is the emphasis placed by Durkheim throughout his monograph on the influence on suicide statistics of demographic factors such as age, gender and marital status.

Durkheim's observation that rates of suicide vary between occupations is supported by contemporary research, although a comprehensive review of secondary sources indicates varying and often opposing hypotheses of relationships between suicide and occupation. A review of the theoretical literature also reveals a tendency for research on occupation-related suicide to focus on a single occupation, or a single industry (and, even then, only a narrow range of these occupations or industries), or a small set of occupations (Boxer et al, 1995; Stack, 2001).

A small number of studies (including those listed in Table 1) have investigated the suicide risk of more than one occupation or category of occupations. But a variety of factors, including those listed below, tend to limit the usefulness of the findings (Stack, 2001):

i) Certain *specific* occupations (e.g. hairdresser)(Lampert, Bourque & Kraus, 1984) are referred to in only one of the articles, which makes it difficult to assess the reliability of the estimate;
ii) Many of the studies do not address female suicide risk (e.g. Lalli & Turner, 1968), a factor which may bias the estimates of suicide risk;
iii) Most of the studies address *general* categories, which makes it difficult to assess the risk associated with specific occupations.

Table 1: Example of American studies that evaluate suicide risk of more than one occupation or category of occupations

Study	Number of occupations or occupational categories investigated
Lampert, Bourque & Kraus, 1984	11 *general* categories (e.g. professional technical-workers)
Milham, 1983	22 *specific* occupations (e.g. shepherd, hairdresser)
Lalli & Turner, 1968	6 *general* categories, but only for males
Labovitz & Hagedorn, 1971	36 *specific* occupations, but only for males

For only a few occupations have detailed explanations of suicide risk been published (Stack, 2001), for example physicians (e.g. de la Monte & Hutchins, 1984), nurses (e.g. Hawton & Vislisel, 1999), farmers (e.g. Malmberg et al, 1999), police-officers (e.g. Violanti et al, 1996), and artists (e.g. Stack, 1997). The multitude of studies focusing on any one of these occupations makes it feasible to assess the reliability and validity of estimates. But, for occupations (e.g. hairdressers) on which studies are rare, the potential for assessing the reliability and validity of estimates is reduced (Milham, 1983).

Research on suicide risk relating to specific occupations is often characterised by inconsistent findings. A review of studies of police suicide, for example, revealed that, of eighteen studies, eleven report a "high" rate of suicide, three report an "average" rate, and four a "low" rate (Stack & Kelley, 1994). Research on physician suicide is marked by similarly inconsistent findings (Bedeian, 1982; Stack, 1998). Stack (2001) suggests that a primary reason for this phenomenon may be the failure to control for the demographic covariates of occupational status, and it is this suggestion that is central to the development of the research question for this thesis.

As well as revealing inconsistent findings for individual occupations, the theoretical literature on occupation-related suicide reveals a deficit in explanation for the differences in suicide risk between occupations (Stack 2001). A comprehensive review of the literature relating to studies conducted in England and Wales reveals two studies that appear to be unusual in reporting suicide data for multiple

occupations (Kelly et al, 1995; Kelly & Bunting, 1998). The methodology used in these studies allows the suicide risk of the range of occupations included in the study to be ranked on the basis of the Proportional Mortality Ratio computed for each occupation. But, by contrast to the American study that is central to this thesis (Stack, 2001), the model of analysis used in these two local studies does not incorporate controls for the very demographic factors (apart from gender) which the authors explicitly acknowledge as being predictor factors for suicide (e.g. age and marital status). Shortfalls such as these form the basis for Stack's assertion that research on occupation-related suicide has tended to neglect multivariate models of analysis and has typically "not sought to weigh the relative importance of certain correlates of occupational status thought to contribute to suicide risk" (Stack, 2001).

In seeking to address this shortcoming, Stack (2001) proposes a logistic regression model of analysis, with controls introduced for four demographic covariates of occupation: gender, race, age and marital status. This model allows the risk of suicide to be computed independent of these four demographic covariates of occupation.

Using this statistical model, Stack (2001) evaluates the suicide risk associated with thirty two occupations or occupational groups, using mortality data drawn from the national Mortality Detail Files (www.ICPSR), which are held by the Inter-University Consortium for Political and Social Research (ICPSR) in the United States, on behalf of the National Center for Health Statistics (NCHS).

It is an attempt to develop an analogous model and test it using local data that is the focus of this thesis. A more detailed discussion of the theoretical assumptions on which this model is based follows in Chapter 3.

In the following section, Stack's *multicausal heuristic model of occupation-related suicide*, which takes into account the influence of the demographic covariates of occupation on the suicide risk associated with the occupation, is discussed.

3. STACK'S MULTICAUSAL HEURISTIC MODEL OF OCCUPATION-RELATED SUICIDE

Suicide is a complex human behaviour, influenced by both individual-level factors (such as mental health, attitudes towards suicide, and living circumstances) and social factors (such as socioeconomic status and unemployment) (Charlton et al, 1992). The chance of an individual committing suicide depends on a combination of the risk and protective factors that are operating in his or her environment at any particular time (www.DPHA). Having a religious affiliation, young dependent children, effective social support, absence of substance abuse, and absence of depression are amongst the most significant protector factors, while chronic ill health, social isolation, substance abuse, homelessness, a history of deliberate self

harm, and stressful life events are all significant risk factors for suicide (Tonda & Baldessarini, 2001).

The potential range of protector and predictor factors, of which occupation is only one, means that any comprehensive statistical model for predicting suicide would necessarily take a multitude of variables into account. To further complicate any analysis of suicide prediction that includes occupation as a variable, the risk associated with 'occupation' itself has a number of sub-variables. Stack (2001) conceptualises the complex relationship between occupation and suicide in terms of a *multicausal heuristic model of occupation-related suicide*, a model which incorporates four major contributory factors to the suicide risk associated with any particular occupation, namely:

i) Stress internal to the nature of the occupation itself
ii) Pre-existing psychiatric morbidity
iii) Opportunity factors
iv) Demographic factors.

Stack's argument is that, for any one occupation, interplay of these factors may either enhance the risk of suicide, or mitigate against it. Thus, a particular occupation may be characterised by, for example, a high level of stress internal to the nature of the occupation itself. But, if it is populated by a demographic group that has a relatively low suicide risk (such as women who are married), this is likely to mitigate against the suicide risk associated with the stress intrinsic to that occupation.

Stack tests only the demographic dimension of the full model. A series of earlier studies derived indicators of risk for individual occupations: labourers (Stack, 1995), dentists (Stack, 1996), and physicians (Stack, 1998). But the study of particular relevance to this thesis derives comparative statistics of suicide risk for thirty two occupations (Stack, 2001).

A description of Stack's *multicausal heuristic model of occupation-related suicide* follows below. Stack (2001) offers only limited discussion of three of these factors (stress internal to the nature of the occupation itself, pre-existing psychiatric morbidity, and opportunity factors) on the grounds that only the demographic dimension of the full model is to be tested.

3.1 STRESS INTERNAL TO THE NATURE OF THE OCCUPATION ITSELF

Stack does not offer an explicit discussion of his conceptualisation of this dimension of the full model. He merely asserts that stress associated with the nature of an occupation may enhance the suicide risk associated with that occupation, and that *client dependence* and *social isolation* are two possible features of occupations that may contribute to the internal stress associated with an occupation (Stack, 2001).

People in client-dependent occupations have been found to have a higher mean suicide rate than people in non-client-dependent occupations (Labovitz & Hagedorn 1971). Client-dependence is the degree to which practitioners of an occupation depend directly on clients for their livelihood, and is typical of physicians in private practice and self-employed trades-people.

The idea that risk may be related to *social isolation* appears to be supported by the finding that the suicide rate for males in the 15-44 age group is above average in areas of Wales that are predominantly agricultural (Kelly et al, 1995). Other studies however (e.g. Hawton et al, 1998) suggest that it is ease of access to means of committing suicide, such as pesticides and firearms, that is the major contributor to the elevated suicide risk of farmers. The varying explanations for the high suicide risk of farmers illustrates the importance of accounting for multiple factors in any analysis of occupation-related suicide.

The stress internal to the nature of policing does not appear to necessarily increase the risk of suicide in this occupation. Although police-officers are at high risk of being exposed to potentially psychotraumatic experiences, a review of studies of police suicide (Stack & Kelley, 1994) reported varying suicide risk rates for police-officers. This variability suggests the possibility that certain characteristics of this occupation may serve as protective factors, such as psychological debriefing exercises after exposure to traumatic events (Bar et al, 2004). This anomaly once again highlights the interplay of protector and predictor factors in suicide risk evaluation.

3.2 PRE-EXISTING PSYCHIATRIC MORBIDITY

The second dimension of Stack's *multicausal heuristic model of occupation-related suicide* refers to the observation that certain occupations appear to attract recruits who have personality traits that enhance their risk of suicide, quite independently of the occupation they occupy (Stack, 2001). Persons with personality traits that enhance risk for suicide (for example, affective disorders) may seek out occupations that are associated with high suicide rates. Two occupational categories that appear to attract such recruits and which many studies have marked as high-risk occupations are artists and psychiatrists (Wasserman, 1992; Andreason, 1987). This finding suggests that, to the extent that self-selection plays a role in the generation of occupational suicide risk, it may not be attributes of the occupation per se that account for its high suicide risk, but rather the attributes of the recruits who are attracted to the occupation

3.3 OPPORTUNITY FACTORS

The third dimension of Stack's model accounts for the observation that the probability of committing suicide may be related to accessibility to means of committing suicide (Charlton et al 1992; Kelly et al 1995: Kelly & Bunting, 1998) and, in the case of medical personnel, expertise in the use of drugs (Hawton et al, 2000).

Most studies consistently report higher than average rates of suicide for medical and dental practitioners, and many studies report higher than average rates for pharmacists, veterinarians, chemical scientists, farmers and forestry workers (Kelly et al, 1995). These occupations have, in common, potentially easy access to various means of committing suicide, such as firearms, pharmaceuticals, and pesticides. The method of suicide favoured by men in the medical and allied professions, all of which have easy access to drugs and poisons, is poisoning by solid and liquid substances (Kelly et al, 1995) The favoured method of suicide for farmers and forestry workers and the second favoured method of vets (occupations which have relatively easy access to firearms) is shooting (Kelly et al, 1995).

The evidence in favour of attributing a high probability of suicide to ease of access to means should however be interpreted with caution. Despite relative ease of access to poisons and firearms, not all studies show an elevated risk for farmers and farm-workers, and opportunity is clearly not a risk factor in the case of military workers, who have easy access to firearms but, according to some studies, do not have a significantly high suicide risk (Stack & Kelley, 1994).

3.4 DEMOGRAPHIC FACTORS

This dimension of Stack's *multicausal heuristic model of occupation-related suicide* is the one on which Stack's logistic regression model of analysis of suicide risk is based, and the most important for this thesis.

A series of studies which derived indicators of suicide risk for individual occupations (as detailed in the introductory part of Section 3) demonstrated the importance of incorporating demographic controls into the analysis (Stack, 1995; Stack, 1996; Stack, 1998). For example, in the case of labourers, even though earlier studies suggest that the associated suicide risk is significantly higher than that of the general population, once controls for gender, marital status, and other covariates of labour status are introduced into the analysis, the relative risk of suicide for labourers is shown to be the same as for the rest of the working-age population (Stack, 1995). This phenomenon is explained by the fact that this occupational category is occupied by a demographic group that is predominantly male and unmarried, two factors which are known to be predictor factors for suicide and

which, before being controlled for, inflate the suicide risk associated with the occupation (Stack, 1995).

Gender and marital status have been shown, in numerous studies, to be significant factors in evaluating suicide risk. Suicide rates in England and Wales tend to increase with age, and are higher for males than for females of the same age (Adelstein & Mardon, 1975). Males are three to four times more likely to die from suicide than are females (although females are more likely to attempt suicide than are males) (www.CDCP; www:DPHA).

Suicide rates for single, widowed, or divorced men are higher than for married men in England and Wales, which suggests that marriage is a protective factor for men (Bunting & Kelly, 1998). For women, the risk is higher for single women than for married, divorced, or widowed women (Bunting & Kelly, 1998).

Since marital status either predicts or protects from suicide, trends in marital status will influence trends in suicide, and therefore also trends in the relationship between occupation and suicide. Changes in the demographic composition of occupations, such as those resulting from a larger proportion of women entering the medical profession and a larger proportion of men entering nursing, is another factor that may influence occupational trends in suicide (Stack, 2001).

People born abroad appear to have a higher risk of suicide than people born locally, although whether this is a function of the characteristics of people who migrate or a function of migration itself is not clear (Adelstein & Mardon, 1975). Geography also appears to be a contributory factor to suicide risk. In the 1992-1996 period, suicide rates were significantly high for both men and women aged 15-44 in the North West, for men aged 15-44 in Wales, and for women 45 and over in the South East. Rates in Scotland were significantly higher overall than those in the rest of England and Wales (Bunting & Kelly, 1980).

A further demographic factor of significance in analyses of occupation-related suicide is socioeconomic status. Men in Social Classes I and II have been shown to have lower suicide rates than men in class V (Drever & Bunting, 1997). For Social Class I, the excess of suicides is concentrated in age groups over 35, with younger age groups having fewer suicide deaths than expected. For social classes IV and V, all age groups demonstrate an excess of suicide over expected deaths, but this excess is particularly marked in the under 35 age groups (Drever & Bunting, 1997).

This overview of demographic factors that may contribute to the suicide risk associated with any occupation completes the discussion of Stack's *multicausal heuristic model of occupation-related suicide*. Before offering a fuller discussion of Stack's methodology for testing the demographic dimension of this model, a discussion of the influence of suicide trends in the general population on occupation-related suicide trends follows.

4. THE INFLUENCE OF SUICIDE TRENDS IN THE GENERAL POPULATION ON OCCUPATION-RELATED SUICIDE TRENDS

Assessment of the suicide risk associated with a particular occupation has to take into account not only the considerations suggested by the four components of Stack's *multicausal heuristic model of occupation-related suicide* (i.e. stress internal to the nature of the occupation itself, pre-existing psychiatric morbidity, opportunity factors, and demographic factors), but also longitudinal suicide trends in the general population.

Studies of trends in suicide since 1911 demonstrate that, over time, there are appreciable shifts in the rates of suicide for each sex, for each age group, and in methods of suicide (Bulusu & Alderson, 1984). These long term trends enhance the complexity of any analysis of occupation-related suicide and mean that occupational trends in suicide should be examined against the background of trends in the general population.

The influence of long term suicide trends in the general population on the relative risk of suicide for individual occupations is demonstrated by two studies of England and Wales (Kelly et al, 1995; Kelly & Bunting, 1998), in which occupations included in the study are ranked for suicide risk on the basis of the Proportional Mortality Ratio (PMR) computed for each occupation, for two separate periods. This allows changes in suicide risk over time to be demonstrated and evaluated. A more detailed discussion of these studies, which will demonstrate (inter alia) the influence of longitudinal suicide trends on assessment of occupational risk for suicide, is offered in Chapter 3, section 3.

In the following section, four factors which influence the recording of primary suicide data, and which may therefore also influence statistical analyses of occupation-related suicide, are discussed.

5. FACTORS INFLUENCING RECORDING OF PRIMARY SUICIDE DATA

A number of official and administrative factors may potentially influence the recording of primary suicide data. These factors, by influencing suicide statistics, may have a bearing on statistical analyses relating to occupation-related suicide, and include:

i) Conventions for classifying suicide as a cause of death
ii) Mortality coding systems
iii) The effect of the Suicide Act of 1961 on the classification of official suicides
iv) Conventions for classifying occupations.

A discussion of these factors follows.

5.1 CONVENTIONS FOR CLASSIFYING SUICIDE AS A CAUSE OF DEATH

In England and Wales, when a death was violent, when it was sudden or of unknown cause, or where there is reasonable cause to suspect that the death was unnatural, there is a legal requirement for the death to be investigated by the Coroner for the district where the death occurred (Adelstein & Mardon, 1974; Charlton et al, 1992). A number of possible verdicts are available in these cases for the coroner to record on the death certificate (www.DU). For the purposes of the present discussion, the most relevant of these are suicide, accident/misadventure, and open verdict.

The International Statistical Classification of Diseases (ICD) (to be discussed in greater detail in Section 5.2) is an international standard for coding mortality data from death certificates, and incorporates a format for reporting causes of death on the death certificate. The reported cause of death is then translated into medical codes through use of the ICD classification structure (www.NCHS). The coroner's verdict of 'suicide' translates to the ICD category of 'suicide and intentional self harm', and corresponds to diagnostic code of E950-E959 in the Ninth Revision of the International Statistical Classification of Diseases (ICD-9) (WHO, 1977) and to X60-X84 in the Tenth Revision (ICD-10)(WHO, 1992-94). For the purposes of the present discussion, this category of deaths will be referred to as 'recorded' suicides.

A verdict of suicide can only be recorded if there is clear evidence that the death was self-inflicted *and* that the person intended to kill him or herself. If there is any doubt about the intentions of the deceased, the death is recorded as 'accidental' or 'open verdict'.

Deaths recorded as accident or misadventure, and those recorded as open verdicts, are categorised in ICD as 'injury or poisoning of undetermined intent', and are coded as indicated in Table 2. This category of deaths is referred to, for the purposes of the present discussion, as 'undetermined deaths'.

Table 2 Correspondence between coroners' verdicts, International Statistical Classification of Diseases (ICD) diagnostic categories, and International Statistical Classification of Diseases diagnostic codes

Coroner's verdict		ICD category	ICD-9 code (pre 2001)	ICD-10 code (from 2001)
Suicide	= 'recorded suicide'	Suicide and self-inflicted injury	E950-E959	X60-X84
Accident/misadventure Open verdict	= 'undetermined death'	Injury undetermined whether accidentally or purposely inflicted	E980-E989 but excluding E988.8	Y10-Y34 but excluding Y33.9

Note: The ICD category 'injury undetermined whether accidentally or purposely inflicted' will, for the purposes of the present discussion, hereafter be referred to as 'deaths of undetermined cause' or 'undetermined deaths'.

The codes E988.8 and Y33.9 are excluded from Table 2 because, since 1979, this category has been used frequently to accelerate death registration in cases where the coroner adjourns the inquest because someone has been charged with an offence in connection with the death or where the case is under police investigation for any other reason. This has given rise to the practice of excluding this category from the wider definition of 'undetermined deaths'. Nearly all cases coded to this category that are eventually resolved turn out, in any case, to be homicide rather than suicide (Charlton et al, 1992).

The criteria for differentiating 'suicide' from 'undetermined deaths' differs between countries, as do the legal and administrative methods used for deriving suicide statistics (Adelstein & Mardon, 1975). Research studies and published national statistics also vary in whether 'undetermined deaths' are included in the statistics for suicide. Some use figures only for 'recorded' suicides, while others, to arrive at a more broadly defined estimate of the number of suicides, include 'undetermined deaths' along with 'recorded' suicides (www.NHS).

It is likely that a large number of deaths classified as 'accidents' are in fact unproven suicides (Adelstein & Mardon, 1975). One category that is particularly difficult to assess is 'accidental poisoning' and, in particular, 'accidental drug overdose'. In the case of the very young and the very old, such deaths are more likely to be accidental than in the middle years, and the same may be true in the case of falling and drowning (Adelstein & Mardon, 1975).

The International Statistical Classification of Diseases (ICD) has been revised periodically since it was introduced in 1901. The category 'injury undetermined whether accidentally or purposely inflicted' was only introduced in 1968, with the adoption of ICD-8. The majority of 'undetermined deaths' are likely to be suicide (Charlton et al, 1992) and, as a result, studies of trends based on data subsequent to 1968 are likely to define 'suicide' as 'recorded suicides + undetermined deaths'. Studies based on data prior to 1968 cannot, by definition, include 'undetermined deaths' in the definition of suicide, while studies that span a time frame that extends either side of 1968 (e.g. Charlton et al, 1992) may incorporate 'undetermined deaths' into the definition of 'suicide' for some parts of the analysis.

Successive revisions of ICD are detailed in Table 3 (below), together with the respective codings for 'recorded suicides' and 'undetermined deaths'. Since some studies define 'suicide' as 'recorded suicides + undetermined deaths', the codes for this category are also indicated.

As this section has demonstrated, the complexities involved in defining 'suicide' may complicate analyses of suicide that have a longitudinal dimension. This complexity is also a consideration when results of different studies, based on different time frames and different versions of ICD, are being compared.

Table 3 International Classification of Disease (ICD) codes for 'suicide and self-inflicted injury' and 'deaths of undetermined cause', 1901 - present

	ICD version	Coding for suicide & self-inflicted injury 'official suicide'	Coding for deaths of undetermined cause 'undetermined deaths'
1901-1910	ICD-1	1900	_
1911-1920	ICD-2	1550, 1560, 1570, 1580, 1590, 1600, 1610, 1620, 1630	_
1921-1930	ICD-3	4650, 1670, 1680, 1690, 1700, 1710, 1720, 1730, 1740	_
1931-1939	ICD-4	1630, 1640, 1650, 1660, 1670, 1680, 1690, 1799, 1710	_
1940-1949	ICD-5	1631 - 1637, 1641 - 1648	_
1950-1957	ICD-6	E970 - E979, E963	_
1958-1967	ICD-7	E970 - E979, E963	_
1968-1978	ICD-8	E950 - E959	E980 - E989 but excluding E988.8
1979-2000	ICD-9	E950 - E959	E980 - E989 but excluding E988.8
2001-present	ICD-10	X60 - X84	Y10 - Y34 but excluding Y33.9

Source: Charlton & Murphy, 1997

5.2 MORTALITY CODING SYSTEMS

The system used for coding deaths may complicate the recording of primary suicide data, as discussed in the previous section, and may also affect the comparability of studies. The International Statistical Classification of Diseases (ICD), which is published by the World Health Organization (WH0), is designed to promote international standardization for processing mortality data from death certificates and has been revised periodically to keep pace with advances in medical knowledge. ICD-9, for example, which was developed in the 1970s and used in England and Wales from 1979 until it was replaced by ICD-10 in 2001, was not adequate for classifying AIDS.

To date, there have been ten revisions of the ICD. The years for which the consecutive revisions of ICD have been used to classify causes of death in England and Wales and the United States respectively are indicated in Table 4 (below). Information on cause of death has been coded to the Tenth Revision of the International Classification of Diseases (ICD-10) since January 2001 (in Scotland, as from January 2000).

ICD-10 (WHO, 1992-1994), the revision that is currently in use, differs from ICD-9 (WHO, 1977) in several ways: it has alphanumeric rather than numeric categories, some chapters have been re-arranged, some titles changed, conditions have been regrouped, and some changes, although only minor, have been made in respect of coding rules for mortality (www.NCHS, ICD-10).

Table 4 Revisions of ICD and years covered, for United Kingdom and United States

Revision of ICD	Years covered: UK	Years covered: US
ICD-1	1901-10	1900-09
ICD-2	1911-20	1910-20
ICD-3	1921-30	1921-29
ICD-4	1931-39	1930-38
ICD-5	1940-49	1939-48
ICD-6	1950-57	1949-57
ICD-7	1958-67	1958-67
ICD-8	1968-78	1968-78
ICD-9	1979-2000	1979-98
ICD-10	2001-present	1999-present

Sources : Griffiths & Brock, 2003; www.NCHS[2]

To interpret trends in cause-specific mortality between ICD-9 and ICD-10, a process called bridge-coding is used for computing *comparability ratios* (www.ONS[4]). Bridge coding involves coding a sample of death certificates independently to both ICD-9 and ICD-10, and comparing the resulting underlying causes of death. Results for equivalent codes or code groups are then presented as comparability ratios of the number of deaths assigned to the given cause of death in the two revisions. The comparability ratio is simply the ratio of the number of deaths coded to a cause in ICD-10 to the number coded to the equivalent cause in ICD-9. A ratio of 1 indicates that the number of deaths coded to that cause is the same in both revisions. A ratio of 0.5 indicates that half as many deaths have been coded to that cause in ICD-10 as in ICD-9 (www.ONS[4]). Table 5 illustrates the effect of the transition from ICD-9 to ICD-10 for the two categories of death relevant to the present discussion. To evaluate trends over a longer time period, a similar computation would be performed for each ICD transition.

As the comparability ratios in this table demonstrate, the transition from ICD-9 to ICD-10 has not had a significant effect in England and Wales on the recording of mortality data for suicide, and therefore is unlikely to affect analyses of occupation-related suicide statistics.

Table 5 ICD-10 / ICD-9 comparability ratios for men and women for suicide and for undetermined deaths

	ICD-9 code	ICD-10 code	Category	ICD10/ICD9 ratio	95% conf. Interval
Women	E950-E959	X60-X84	Suicide & self-inflicted injury	0.999	(0.996 - 1.001)
	E980-E989	Y10-Y34	Undetermined deaths	1.014	(0.999 - 1.029)
Men	E950-E959	X60-X84	Suicide & self-inflicted injury	1.000	(0.998 - 1.001)
	E980-E989	Y10-Y34	Undetermined deaths	1.010	(1.000 - 1.020)

Source : (www.ONS[4]).

5.3 THE SUICIDE ACT OF 1961

The effect of the Suicide Act of 1961 is the third of the four factors addressed by this section that may potentially affect the recording of suicide data.

Suicide was a criminal act until the promulgation of the Suicide Act in 1961. Abolishing the criminal aspect of suicide might have been expected to affect the classification of 'recorded' suicides and 'deaths of undetermined cause' respectively, and to result in an increased rate of recording of official suicides and a decreased rate of recording of deaths of undetermined cause. The rate of recorded suicide did indeed rise over the two years following the promulgation of the Suicide Act. But, paradoxically, so did the rate for poisoning (accidental and open verdict cases) – and at an even greater rate than did the rate for 'official' suicide (Adelstein & Mardon 1975). This trend continued for two years, although the rates for both 'suicide' and 'poisoning' (which included both accidental and open verdict cases), having peaked in 1963, fell consistently thereafter until 1974.

The practice of recording 'open verdicts' other than from 'poisoning' was not adopted until 1967 (Adelstein & Mardon, 1975). This means it is difficult to evaluate the effect of the Suicide Act on the recording of official (i.e. recorded) suicides and deaths from 'undetermined' causes respectively for the period immediately following the promulgation of the Suicide Act in 1961, until 1967. For this period, rates for 'estimated' suicides (that is, the sum of official suicides and poisoning and other open verdicts) cannot be computed because of the absence of statistics for 'other open verdicts'. It is only with effect from 1967, with adoption of the practice of recording 'open verdicts' other than from 'poisoning', that statistics for 'estimated' suicides (i.e. official suicides + poisoning + other open verdicts) can be computed, as illustrated in Table 6.

Table 6 Ratio of official suicides to estimated suicides for persons of all ages, 1967-1974

(Suicide rates per million)	1961	1962	1963	1964	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974
Official suicides	111	118	120	116	107	103	97	95	89	81	81	77	78	79
Poisoning (*accidental & open verdict cases*)	39	39	44	37	35	37	33	32	33	32	32	29	29	28
Other open verdicts							11	13	14	11	10	11	11	12
Estimated suicides (*official suicides + poisoning + other open verdicts*)							142	139	137	124	123	116	117	119
Ratio of official suicides to estimated suicides							.68	.68	.65	.65	.66	.66	.67	.66

Source : Adelstein & Mardon 1975: 14, Table 1.

According to one study (Adelstein and Mardon, 1975), even though the rates for official suicides and poisoning respectively declined steadily after 1963, the rates for 'open verdicts' other than poisoning tended to be fairly consistent from the time they were first recorded separately in 1967. When rates for 'estimated' suicides are computed (a value which is equivalent to the sum of official suicides, poisoning, and other open verdicts), the overall effect is that the ratio of 'official' suicides to 'estimated' suicides was fairly constant (65-69%) for the period 1967-1974, as indicated Table 6. This suggests that deaths from undetermined causes (which are represented, for the purposes of the present discussion, by the sum of 'poisoning' and 'other open verdicts') varied at a similar rate to that of 'official' suicides for the period 1967-1974. The implication of this finding is that the 1961 Suicide Act does not appear to have markedly affected the classification of these categories of death, and therefore is unlikely to have any effect on analyses of occupation-related suicide trends.

5.4 CLASSIFICATION OF OCCUPATIONS

This section offers only a brief discussion of the influence that conventions for classifying occupations have on suicide statistics, and thus on statistical analyses of occupation-related suicide risk. Conventions for classifying occupations will be discussed more extensively in Chapter 5. But for the purposes of the present section, it is relevant to note that the use of different protocols for classifying occupations may potentially influence occupational analysis of suicide data. This consideration is particularly relevant in the case of comparative studies, where different occupational classification systems may have been used, and also in the case of longitudinal studies, where transitions between revised versions of a single occupational classification system may produce effects similar to those produced by transitions between revisions of ICD, as discussed in an earlier section.

6. CONCLUSION

This chapter started with an overview of theoretical literature relating to a study of occupation-related suicide, and highlighted the ongoing relevance of Durkheim's theorising on suicide to contemporary studies of occupation and suicide. From a discussion of the literature emerged a series of theoretical considerations that are relevant to any analysis of occupation-related suicide risk. This included identification of four factors that could potentially influence the recording of primary suicide data and thus have a bearing on analyses of occupation related suicide, i.e. conventions for classifying suicide as a cause of death, mortality

coding systems, the effect of the Suicide Act of 1961 on the classification of official suicides, and conventions for classifying occupations.

As this chapter has demonstrated, conventions for defining 'suicide' are complex and differ between studies, between countries, and between historical eras. The complexities involved in defining 'suicide' tend to complicate an analysis of occupation-related suicide. And the introduction of the 'undetermined' category of deaths with ICD-9 in 1968 further complicates analyses that have a longitudinal dimension. Many - if not most - of the adult deaths recorded as open verdicts are likely to have been suicides. But, even though the criteria for suicide (i.e. the death was self-inflicted and the person intended to kill him or herself) may have prevailed, if they were not proven, then the death is recorded as an open verdict instead. Thus it is probable that there is some under-recording of suicide (Charlton et al, 1992). But, as this chapter has demonstrated, the ratio between official (or recorded) suicide rates and estimated suicide rates (i.e. official suicides + undetermined deaths) was fairly constant for the period 1967-1974, which suggests that, even though there may be some under-recording of official suicides, this would not affect analyses based on 'estimated' suicides.

The definition of 'suicide' used in England and Wales varies between studies. Some adopt a definition of suicide as 'official (or 'recorded') suicide + undetermined deaths' (i.e. 'estimated' suicides) (e.g. Kelly et al, 1995). Others, particularly those that present data for time periods prior to 1968, take into account only 'official' suicides (e.g. Bulusu & Alderson, 1984), while studies that take into account statistics from both sides of 1967 are likely to use two definitions of 'suicide': 'recorded suicides' and 'suicides + undetermined deaths' (e.g. Charlton et al, 1992).

This chapter has also shown that the Suicide Act of 1961, which abolished the criminal aspect of suicide, did not appear to markedly affect the classification of suicide deaths. This finding is significant in the case of suicide analyses that have a longitudinal dimension.

Transitions between consecutive revisions of mortality coding systems have been shown in this chapter to have a potential influence on the comparability of mortality data that is based on different revisions of ICD. This is particularly relevant in the case of longitudinal studies. The tenth revision of ICD (ICD-10) incorporates only minor changes in respect of coding rules for mortality (www NCHS[3]). Even so, for some causes of death, changes in the rules for selecting the underlying cause of death means that mortality statistics will demonstrate a discontinuity in trends (Rooney & Smith, 2000). However, for the two coding categories relevant to suicide analyses, the effect of the transition from ICD-9 to ICD-10 is not significant. The effect of the transition from ICD-7 to ICD-8 in 1968 has a more significant effect than that of the transitions between ICD-8, ICD-9, and ICD-10. This is because of the complication introduced by the introduction of the 'undetermined deaths' category in ICD-8.

As the discussion of these four factors that could potentially influence analyses of occupation-related suicide has shown, their effect is particularly relevant to studies with a longitudinal dimension which spans two or more revisions of mortality coding frames and occupational coding frames. These factors also influence comparability of results from different studies in cases where the studies have employed different versions or revisions of mortality or occupational coding frames.

This chapter has also provided a brief discussion of the predictor and protective factors for suicide that emerge from an overview of the literature, with occupation being only one of many predictor factors. Stack (2001) conceptualizes the complex relationship between occupation and suicide in terms of a *multicausal heuristic model of occupation-related suicide*, a model which incorporates four major contributory factors to the suicide risk associated with any one occupation, i.e. stress internal to the nature of the occupation itself, pre-existing psychiatric morbidity, opportunity factors, and demographic factors. In a series of studies of individual occupations and one comparative study of 32 occupations, Stack tests only the demographic dimension of the full model. The methodology used in these studies is based on a logistic regression model of analysis, incorporating controls for a series of demographic covariates of occupation. It is this methodology, and an attempt to replicate it using data from England and Wales, that forms a central theme to this thesis. A more extensive discussion of this methodology will be offered in Chapter 3.

A thorough review of the theoretical literature, together with discussion I had with staff at the Office for National Statistics, suggests that the methodology used by Stack has not to date been applied in studies of occupation-related suicide in England and Wales. Studies of England and Wales (Kelly et al, 1995; Kelly & Bunting 1998) offer a ranking of the suicide risk of a range of occupations on the basis of the Proportional Mortality Ratio for each occupation. But the model of analysis that these studies employ does not incorporate controls for the very demographic factors (apart from gender) which the authors acknowledge to be predictor factors for suicide, such as age and marital status. This apparent shortfall, together with the evidence that the methodology used in Stack's American studies has not been tested using data from England and Wales, suggests a gap in the literature.

In the next chapter I will discuss Stack's methodology and findings in detail and, in Chapter 4, will formulate a research question for this thesis with a view to attempting to replicate Stack's methodology for multivariate analysis of occupation-related suicide, using local data.

CHAPTER 3

EVALUATION OF THREE STUDIES RELEVANT TO THE RESEARCH QUESTION

1. INTRODUCTION

As was discussed in Chapter 2, measures of suicide risk are precarious over time because of longitudinal trends that influence suicide rates. These trends are influenced by phenomena such as changes in the demographic composition of the population and changes in the availability of suicide methods. For example, a decrease in the number of prescriptions for barbiturates in England and Wales between 1966 and 1973 was accompanied by a 25% decrease in the rate for suicides by solid and liquid substances (Adelstein & Mardon, 1975). The precariousness of measures of suicide risk in the general population is especially relevant in the case of analyses of occupation-related suicide risk that have a longitudinal dimension. Analyses of occupation-related suicide risk are further complicated by the complexities involved in categorising occupations and in coding information from death certificates, as was discussed in Chapter 2.

Even for comparative studies, measures of suicide risk must be a tentative evaluation only, since the reliability of estimates is difficult to establish because of influences such as social class factors. As will be discussed in the present chapter, in the case of studies that employ the Provisional Mortality Ratio (PMR) as a measure of occupational suicide risk, there is the possibility that, for Social Class I and II occupations, the PMR may be a better measure of the risk of *not* dying from other causes than the risk of dying from suicide (Kelly & Bunting, 1992). The influence of socioeconomic factors is shown to be a significant consideration when evaluating and comparing the results of the two UK studies (Kelly et al; Bunting and Kelly) and the American study (Stack, 2001) that are central to discussion in this chapter.

As Chapter 3 will show, the number of occupations found to be at either significantly high or significantly low risk of suicide, on the basis of an analysis of Provisional Mortality Ratios, is far higher in two studies in England and Wales (Kelly et al, 1995; Kelly & Bunting, 1998) than the eight identified in the American study (Stack, 2001) in which multivariate logistic regression odds ratios are used for evaluating occupational risk for suicide.

In section 2 of this chapter, a comprehensive discussion is offered of the theoretical assumptions and methodology on which Stack's study is based. The results of the study will be discussed, with specific reference to the findings for occupations indicated to be at significantly *high* (as opposed to significantly *low*) risk for suicide. In section 3, the first of the two relevant studies of suicide in England and Wales is described (Kelly et al, 1995), and the theoretical assumptions on which use of the Proportional Mortality Ratio for suicide is based are discussed. Discussion of this study also highlights the influence of longitudinal suicide trends on measures of occupation-related suicide risk by comparing findings for one time period with those of another.

Discussion of the second study (Kelly & Bunting, 1998) highlights the importance of controlling for the demographic covariates of occupation, a consideration which is central to the theorising outlined by and methodology employed by Stack. This section also examines the overlap between the findings of Kelly & Bunting's and Stack's studies, and demonstrates inconsistencies that arise as a result of different theoretical assumptions.

The final section summarises the different theoretical assumptions and methodologies on which the studies are based, and highlights the complexities involved in comparing the findings of studies that use different indicators for measuring suicide risk. The discussion also highlights an area of theoretical interest that provides the focus for development of a research problem for this thesis: an investigation of occupation-related suicide risk in England and Wales, based on a multivariate model of analysis analogous to that employed by Stack

2. STACK'S U.S. STUDY: METHODOLOGY AND FINDINGS

As was discussed in Chapter 2, research on suicide risk relating to specific occupations is often characterised by inconsistent findings, a phenomenon which may at least partly be explained by the failure to control for the demographic covariates of occupational status (Stack 2001). In seeking to address this shortcoming, Stack (1995, 1996, 1998, 2001) proposes a logistic regression model of analysis of occupational suicide risk, with controls introduced for four demographic covariates of occupation: gender, race, age and marital status. This model allows the risk of suicide to be computed independent of these four demographic covariates of occupation.

Logistic regression is a predictive analysis, based on binomial probability theory, that enables the probability of a dichotomous dependent variable (in this case, 'suicide') to be computed. This method of analysis can incorporate more than one predictor variable, and can therefore be used to evaluate the impact on the dependent variable of several predictor variables (in this case, gender, race, age and

marital status) at the same time. The probability of the outcome (in this case, 'suicide') is represented by 'odds' ratios.

Using this statistical model, Stack (2001) evaluates the suicide risk associated with 32 occupations or occupational groups (although results are reported for only 30 of these), using mortality data drawn from the national Mortality Detail Files (www.ICPSR), which are held by the Inter-University Consortium for Political and Social Research (ICPSR) in the United States, on behalf of the National Center for Health Statistics (NCHS).

The first stage of the analysis is based on a bivariate logistic regression model. A dichotomous predictor variable is derived by comparing each of the 32 occupational groups with 'all other occupations'. The dependent variable is (by definition) a binary variable, where 1=death from suicide and 0=death from all other causes. In this way, the bivariate logistic regression odds ratio for each of the 32 occupational groups is computed, using the logistic regression utility programs in SAS (Statistical Analysis System).

In the second stage of the analysis, demographic controls are introduced into the logistic regression equation. The predictor variables and their associated codes are shown in Table 7. Stack justifies the dichotomisation of 'race' to 'whites' and 'all others' on the grounds that whites have been shown in other studies to have a higher suicide rate than other racial groups (Stack, 2001). The dichotomisation of 'marital status' to 'married' and 'all others' is justified by findings in many other studies that 'married' status is a protective factor against suicide (Stack, 2001; Durkheim, 1952). Further discussion about marital status as a predictor or protector factor for suicide is offered later in this thesis. 'Age' is a continuous variable. The age range for this analysis is defined as 18-64, which represents (for the purposes of this study) the working-age population (Stack, 2001).

Table 7 Demograghic covariates and associated codes incorporated into Stack's multivariate logistic regression equation

Variable	Nature of variable	Codes	
Gender	Binary	1 = male,	0 = female
Race	Binary	1 = white,	0 = all others
Marital status	Binary	1 = married,	0 = all others
Age	Continuous (in years), 18-64		

Source: Stack 2001

The definition of occupational groups for this study was based on standard federal occupational codes (Stack, 200). The 32 occupational groups included in the analysis (see Table 8) include representative occupations from six major occupational headings and do not constitute the universe of all occupations (Stack, 2001). They were selected principally on the basis of representing at least 50 suicides. In some instances, such as for 'managers' and 'executives', occupational

groups were combined in order for sufficient suicides to be represented, since the larger the number of cases in the cell, the greater the validity of the analysis.

The study is based on "official suicide statistics" (Stack 2001: 389), but it is not clear from the report whether 'official suicide statistics' refers only to 'recorded' suicides, or whether 'undetermined deaths' are included in this categorisation. The lack of clarity about the definition of suicide has implications for the comparability of results with those from other studies, as will be discussed later in this chapter.

The study is based on 9,499 suicides and 134,386 deaths from all other causes for the year 1990. Data for 1990 are drawn from the National Mortality Detail File (www.ICPSR[1]) in respect of suicide data from the twenty one states that report occupational data in the National Mortality Detail File (Stack, 2001). A more detailed discussion of the National Mortality Detail File will be offered in Chapter 6, and will highlight the legal basis for the mandate that the data held in the National Mortality Detail File be made publicly available for secondary analysis (www.HHS[2]). As the discussion will show, this mandate contrasts directly with the situation in England and Wales, where legislation precludes access to individual-level census and vital-events data by parties external to the Office for National Statistics.

Cause of death is coded for this study (as it is for the local studies that will be discussed in the following section) using the eighth and ninth revisions of the International Statistical Classification of Diseases (ICD-9 and ICD-10).

Table 8 Occupational groups in Stack's study

Major occupational heading	Representative occupations	
Managerial/professional	Accountants Artists Dentist Doctor Elementary school teachers Engineers	Executives/managers Lawyers Mathematicians and scientists Nurses Professors Social workers
Clerical	Bookkeepers Clerks Postal workers	
Service	Bartenders Cooks Police-officers Private security	
Agricultural and extractive	Farmers Farm workers Miners	
Skilled mechanical	Auto mechanics Carpenters Electricians Machinists Plumbers	
Semi-skilled and unskilled manual	Labourers Truck-drivers (heavy equipment) Welders	

Source: Stack, 2001

A comparison of the bivariate regression odds ratios and the multivariate logistic regression odds ratios for the 32 occupations included in the study shows that 15 occupations are indicated to be at either significantly high or significantly low risk for suicide on the basis of bivariate logistic regression odds ratios; but, after demographic controls are introduced into the analysis, only 8 occupations are indicated to be at either significantly high or significantly low risk. For the purposes of this thesis, results will be discussed only for those occupations associated with a significantly *high* risk of suicide on either bivariate or multivariate analysis (as indicated in Table 9).

Of the six occupations shown to be at high risk of suicide after demographic controls are introduced (i.e. doctor, dentist, mathematician & scientist, artists, nurses, social workers), four (i.e. doctor, dentist, mathematician & scientist, artists) were indicated to be at high risk even before demographic controls were introduced (see Table 9). Two of the six occupations (i.e. nurses and social workers) were not indicated to be at significantly high risk on the basis of bivariate analysis, but introducing demographic controls had the effect of lifting the occupations into high-risk status. Eight occupations or occupational groups were indicated to be at high risk on the basis of bivariate analysis, but introducing demographic controls mitigated against this effect.

Table 9 Comparative bivariate and multivariate logistic regression odds ratios for 14 occupations investigated by Stack (2001) to show effect of introducing demographic controls into logistic regression analysis

			Bivariate Logistic Regr Odds Ratio	Multivariate Logistic Regr Odds Ratio
A	4 occupations at significantly high risk both before and after demographic controls introduced	Doctor	1.94*	2.31*
		Dentist	4.45*	5.43*
		Mathematicians & scientists	1.85*	1.47*
		Artists	2.12*	1.30*
B	2 occupations at significantly high risk only after demographic controls introduced	Nurses	0.99	1.58*
		Social workers	1.41	1.52*
C	8 occupations at significantly high risk only before demographic controls introduced	Miners	1.33*	1.09
		Machinists	1.63*	1.23
		Auto mechanics	1.41*	0.87
		Electricians	1.32*	0.99
		Plumbers	1.63*	1.23
		Carpenters	2.00*	1.22
		Welders	1.46*	1.01
		Labourers	1.31*	0.93

Source: Stack, 2001 *** p < 0.05 for the associated logistic regression coefficient**

Note: the denotations A, B and C represent cross references to the legend for Table 12.

Stack argues that these findings (which will be discussed in more detail in later sections of this chapter) underscore the importance of incorporating demographic controls in the assessment of occupational risk for suicide, and provide "the first systematic evidence on the problem [of failing to incorporate demographic controls in analysis of occupation-related suicide risk] for the United States" (Stack 2001: 386).

In the following section, key features of the two local studies which are relevant to the present discussion (Kelly et al, 1995; Kelly & Bunting) will be discussed, and the findings contrasted with those of Stack's American study.

3. TWO STUDIES IN ENGLAND AND WALES: METHODOLOGY AND FINDINGS

As discussed briefly in Chapter 2, two studies are unusual in reporting suicide data for *multiple* occupations in England and Wales (Kelly et al, 1995; Kelly & Bunting, 1998). The methodology on which these studies are based allows the suicide risk of the occupations included in the study to be ranked on the basis of the Proportional Mortality Ratio (PMR) computed for each occupation. The Proportional Mortality Ratio for suicide is a measure of the rate of suicide within a specific population group (in this case, an occupation) relative to all other causes of mortality, and enables the effect of suicide on that occupational population to be examined. The PMR is computed as follows:

$$\text{PMR(suicide)} = \frac{\text{Observed deaths from suicide}}{\text{Expected deaths from suicide}} \times 100$$

'Expected deaths' are computed by applying the proportion of total deaths due to suicide in the comparison population (which is, for the purposes of the study being discussed, men aged 20-64 or women aged 20-59 in the general population) to the total number of deaths in the occupation of interest (Kelly et al, 1995).

A PMR that is greater than 100 indicates that, for that gender, the given occupation has a greater proportion of deaths from suicide by comparison with that age-group and gender in the general population. A PMR of less than 100 indicates a *lower* proportion of deaths from suicide.

Caution should be exercised when interpreting PMRs, since a high PMR for suicide *may* represent a true risk, but it may also represent a deficit in deaths from other causes. Thus an occupation might reflect a high PMR for suicide, even if the suicide rate for that occupation is lower than the national rate, if it is the case that mortality from 'all causes' in that occupation is low (Kelly & Bunting 1998).

Conversely, an occupation may reflect a low PMR for suicide even if the suicide rate for that occupation is higher than the national rate, if it is the case that mortality from 'all causes' in that occupation is high. This phenomenon is particularly relevant in the context of evaluating the relative suicide risk of various occupations. Many of the occupations that research has consistently indicated to be at high risk, for example dentists (e.g. Stack, 1996) and doctors (e.g. Hawton et al, 2000), are classified as higher social status occupations (Lampert et al, 1984). But many studies have suggested that, in the general population, Social Classes I and II have a lower rate of mortality from 'all causes' than other social classes (Kelly et al, 1998). Thus a high PMR for suicide for Social Class I and II occupations may say less about the suicide risk for these occupations than about the overall rate of mortality for Social Classes I and II (Kelly et al, 1998). A converse caveat is likely to apply to Social Class IV and V occupations: since the overall mortality rate in the general population is higher for Social Classes IV and V, a low PMR for suicide for Social Class IV and V occupations may not be a true reflection of the suicide risk associated with those occupations.

The earlier of the two studies referred to in this section (Kelly et al, 1995) examines the effect of occupation on suicide mortality by calculating PMRs for the period 1982-1992 for ten occupations for men and ten for women that were identified in other studies (Charlton et al, 1992; Drever, 1995) to be at highest risk of suicide.

The data period 1982-1992 was split into two, 1982-1987 and 1988-1992, thus introducing a longitudinal dimension to the analysis and enabling a comparison of suicide ratios over time for each of the occupations included in the analysis. Occupations are ranked on the basis of their respective PMRs for suicide for the *first* time period. The definition of suicide used for this study is 'suicides + undetermined deaths'. The corresponding ICD-codes are shown in Table 3.

Analyses were restricted to the 16-64 age-group for men and 16-59 age-group for women, with the upper limit reflecting retirement age. The analysis was based on 34162 male suicides (with the distribution of 17549 to the earlier time period and 16613 to the later period) and 11187 female suicides (6536 in the earlier time period and 4651 in the later period). The source of primary data is not specified but, as will be discussed in detail in Chapter 6, the source in England and Wales of mortality data derived from death certificates is necessarily the Office for National Statistics.

Cause of death is coded using the Ninth Revision of the International Classification of Diseases (ICD-9). The protocol used for defining occupations and occupational groups is not specified, although the classification of occupations is likely to have been based on the Classification of Occupations 1980 (CO80), which was used in England and Wales from 1980, and Standard Occupational Classifications (SCO90), which replaced CO80 in 1990. A more detailed discussion about conventions for classifying occupations will be offered in Chapter 5.

The occupations included in this analysis are shown in Table 10. The table also indicates the direction of change of the PMR for each occupation between the two time periods.

As Table 10 demonstrates, for men, the PMR for seven of the occupations decreased between the two time periods, including that for dentists and doctors. Research has consistently shown the rate of suicide amongst physicians and dentists to be high, which makes the fall in the PMR for these professions worthy of note. As pointed out earlier in this section, a fall in the PMR for an occupation does not necessarily indicate a decrease in the rate of suicide for that occupation but might indicate that the rate of deaths from other causes is increasing. The example shown by this study of a decrease in PMR for medicine and dentistry, occupations that have consistently been shown to be at high risk of suicide, illustrates the need to interpret findings based on PMRs with caution.

Table 10 Ten highest risk occupations for men aged 16-64 and women aged 26-59 for 1982-1992, ranked according to PMR for suicide for 1982-1987

	Men		Women
+	Vets	-	Government inspectors
- -	Librarian, information officer	-	Vets
-	Dental practitioners	-	Medical practitioners
-	Farmers	-	Pharmacists
-	Pharmacists	- -	Physiotherapists
-	Forestry workers	+	Ambulance-women
-	Medical practitioners	- -	Health professionals not elsewhere classified
-	Hotel porters	-	Literary and artistic professionals
+	Chemical scientists & engineers	- -	Teachers in higher education
+	University academic staff	-	Nurses

Source: Kelly et al, 1995

\+ PMR increased between 1982-87 and 1988-92
\- PMR decreased between 1982-87 and 1988-92 but remained above 100
\- - PMR decreased between 1982-87 and 1988-92 and fell to below 100

Despite the caveat detailed in the previous paragraph, the findings of this study are of particular significance in demonstrating the precariousness of suicide statistics over time, a phenomenon which was referred to in Chapter 2. This precariousness is a factor that must be taken into account particularly in the case of comparative studies, and will be discussed in more detail later in this chapter.

Of further note is that, of the ten high-risk male occupations identified for this study, all but two (i.e. hotel porters and forestry workers) are Social Class I or II occupations. The ten female occupations are also marked by a preponderance of Social Class I and II occupations. As suggested earlier in this section, this observation must be interpreted with caution before an assumption is made that higher rates of suicide risk are associated with Social Class I and II occupations. The high PMR for these occupations may say more about the risk of *not* dying

from other causes than about the risk of dying from suicide (Kelly & Bunting, 1992). This is by contrast with Social Class VI and V occupations, where a low PMR for suicide may be a more meaningful indicator of the risk of dying from other causes, rather than the risk of dying from suicide. Of note in this context is the finding by Stack (as discussed in the previous section) that eight occupations of lower socioeconomic status appeared to be at low risk of suicide on the basis of bivariate analysis. But introducing demographic controls into the logistic regression analysis confounded the assignment of *low* suicide-risk to these occupations and, on the basis of multivariate analysis, each was indicated to be at high risk.

In the second of the two studies that are the focus of this section (Kelly & Bunting, 1998), occupations to be investigated were selected on the basis of the number of deaths occurring in the 1982-1987 time period to men and women in the 20-74 age-group. For men, occupations that represent 20 or more suicide deaths were selected and, for women (whose suicide rate is lower than that for men), occupations with 10 or more deaths were selected. Further occupations, which had lower numbers of deaths but have been shown in earlier studies (Kelly et al, 1995) to have high PMRs, were also included (e.g. male and female veterinarians).

Although selection of *occupations* was based on examination of the 20-74 age-group, selection of *subjects* for the study was based on a narrower age-range. Analyses were restricted to the 20-64 age-group for men and 20-59 age-group for women. The lower limit was set to 20 rather than 16 because more than 40 percent of men aged 16-19 are classified as unoccupied (Kelly & Bunting, 1998; Drever, 1997).

As in the case of the study discussed earlier in this section (Kelly et al, 1995), the protocol used for defining occupations and occupational groups is not specified, although, as before, definitions are likely to have been based on CO80 and SCO90. The source of primary data for the analysis is also not specified (as in the case of the study discussed earlier), but data will almost certainly have been drawn from the databank held by the Office for National Statistics, a topic which will be discussed in greater detail in Chapter 6. Cause of death is coded using the Ninth Revision of the International Classification of Diseases (ICD-9). Suicide deaths are defined as 'suicides and deaths from injury and poisoning undetermined whether accidentally or purposely inflicted', which corresponds to codes E950-E959 and E980-E989, excluding E988.8 (see Table 3).

In common with the study discussed earlier in this section, the more recent study introduces a longitudinal dimension to the analysis by presenting data for two time periods. The earlier study derives the two time periods by splitting the overall time period 1982-1992 represented by the study. By contrast, the more recent study defines two independent time periods (1982-1987 and 1991-1996) and analyses relevant primary data independently for each time period. Thus

occupations which are identified to be at high risk in the earlier time period are not necessarily represented in the later time period, and vice versa.

Analyses were performed for 139 occupations for men and 47 for women, for two time periods: 1982-1987 and 1991-1996. For each time period, Proportional Mortality Ratios (PMR) are presented for occupations whose ratios are significantly higher or lower than for the general population of England and Wales for the relevant gender and age-group for the relevant time period (i.e. where the 95% confidence interval excluded 100).

For the earlier time period, 17 occupations or occupational categories were found to have significantly high PMRs for men, and 15 in the later time period (as indicated in Table 11). Only 7 occupations were common to both periods. For women, 9 occupations or occupational categories were found to have a significantly high PMR in each time period. Of these, only 5 categories were common to both time periods. This phenomenon (as illustrated in Table 11) once again demonstrates the precariousness over time of estimates for suicide risk.

By demonstrating different trends for men and women, these results emphasise the significance of controlling for gender in analyses of occupational risk for suicide. Even though the authors acknowledge that age and marital status are predictor factors for suicide (Kelly et al 1995; Kelly & Bunting, 1998), gender is the only demographic variable controlled for in the two studies discussed in this section. But the effect of controlling for it highlights the importance of controlling for other demographic covariates of occupation, a consideration which is central to the theorising outlined by and methodology employed by Stack (as discussed in Section 2).

Table 11 Numbers of occupations with significantly high PMR for suicide for men and women for 1982-1996, 1982-1987 and 1991-1996 and occupations common to time-periods 1982-1987 and 1991-1996

	1982-1996	1982-1987	1991-1996	Occupations common to both time-periods	
Men	32	17	15	7	Farmers, horticulturists, farm managers Sales reps for property and services Shop salesmen and assistants Gardeners, groundsmen Dental practitioners Medical practitioners Pharmacists
Women	19	9	9	5	Domestic housekeepers Medical practitioners Nurses and nurse-administrators Students Waitresses

Source : Kelly & Bunting, 1998: 35, table 5

Examining the overlap between the findings of Kelly & Bunting's study and Stack's study provides insights into how the differing methodologies yield some results that are consistent between the studies and others that are discrepant. Table 12 represents those occupations that are included in both studies *and* reflect statistically significant results for risk, a total of only 6 occupations.

Table 12 Illustration of overlap between the findings of Kelly & Bunting and Stack studies

Kelly & Bunting (1982-87 and 1991-1996)					Stack (1992)			
	Men		Women		Men & women			
	82-87	91-96	82-87	91-96		A	B	C
Medical practitioners	■	■	■	■	Doctor	■	■	■
Dental practitioners	■	■			Dentist	■	■	■
Nurse-administrators, nurses			■	■	Nurses		■	■
Labourers & unskilled workers	■				Labourers			■
Plumbing, heating, gas fitting	■				Plumbers			■
Carpenters and joiners		■			Carpenters			■

Source: Kelly & Bunting, 1998; Stack, 2001.

Legend:

Kelly & Bunting study	
Occupations in Kelly & Bunting study that overlap with occupations from Stack study	
	Occupations at significantly high risk of suicide
	Occupations at significantly low risk of suicide
Stack study	
Occupations in Stack study that overlap with occupations from Kelly & Bunting study	
A	Occupations at significantly high risk both before & after demographic controls introduced
B	Occupations at significantly high risk only after demographic controls introduced
C	Occupations at significantly high risk only before demographic controls introduced

Note: the denotations A, B and C represent cross references to the legend for Table 9

Stack's findings for doctors, as indicated in Table 12, suggest that doctors are at high risk of suicide, both before and after demographic controls are introduced into the analysis. Introducing demographic controls into the logistic regression analysis does not appear to have counteracted the risk attributed to this occupation on the basis of bivariate analysis. This finding suggests that the suicide risk associated with this occupation may be independent of the influence of the demographic factors that are incorporated into the multivariate analysis. This observation is supported by the finding of Kelly & Bunting's study that this is a high-risk occupation for both men and women. The gender consistency demonstrated by Kelly & Bunting can, however, be only a partial explanation for the effect demonstrated by Stack (i.e. that introducing demographic controls into the logistic regression analysis does not counteract the suicide risk associated with this occupation on the basis of bivariate analysis) since Kelly & Bunting's findings account only for gender, and not for any of the other demographic controls incorporated in Stack's multivariate analysis (i.e. age, marital status, and race).

Kelly & Bunting's findings for doctors also demonstrate consistency in risk, for both men and women, between the two time periods included in the analysis. This longitudinal consistency suggests that the suicide risk associated with this occupation is not precarious. In light of Stack's finding that demographic factors do not appear to any significant degree to account for this risk, it is possible that it could be better accounted for in terms of other dimensions of Stack's *multicausal heuristic model of occupation-related suicide* (as discussed in Chapter 2, section 3). Of relevance to the 'opportunity factors' component of this model is the evidence offered by many studies (e.g. Hawton et al, 2000) that the relative ease of availability of pharmaceuticals and expertise in their use may contribute to the elevated risk of suicide in medical practitioners. But the 'stress internal to the nature of the occupation itself' component of Stack's model may be of equal significance. This thesis does not attempt to explore this dimension of Stack's model further but, as will be discussed in Chapter 6, this component of the model warrants further exploration.

Stack's results associate a high level of suicide risk with dentists, both before and after demographic controls are introduced into the analysis. But the findings for dentists contrast with those for doctors in that Kelly & Bunting's study offers findings for male dentists only. Female dentists were excluded from this study because this is a statistically infrequent role-set which attracts few recruits and therefore did not meet the criterion for inclusion in the study (i.e. 10 or more deaths for the relevant time period) (Kelly & Bunting, 1998). The absence of suicide-risk information for female dentists (as well as absence of information relating to demographic factors other than gender) reduces the potential for comparing Kelly & Stack's findings for dentists with those of Stack's.

As in the case of doctors, Stack's findings for dentists suggest that factors other than demographic ones may be at play in explaining the elevated risk of suicide for dentists. By contrast, Stack finds that, for nurses, demographic influences are significant in explaining the risk of suicide associated with this occupation. The suicide risk of nurses was shown to be *not* significant on the basis of bivariate analysis but, after demographic controls were introduced into the analysis, a significantly high risk for suicide was demonstrated. This effect (i.e. shifts in the risk associated with the occupation once demographic controls are introduced) illustrates the importance of introducing demographic controls into the analysis, particularly in the case of an occupation that is traditionally occupied by women, who are at lower risk of suicide than men. Stack's finding (i.e. that the suicide risk associated with nursing is significantly high, but only after demographic controls are introduced into the analysis) is consistent (in this case) with the findings of Kelly & Bunting.

Two categories of occupation (i.e. labourers & unskilled workers; plumbing, heating, gas-fitting) are shown to be associated with a significantly *low* risk of suicide in Kelly & Bunting's study for the first time period. By contrast, Stack finds

that these occupational groups are associated with a significantly *high* suicide risk, but only on the basis of bivariate analysis. Once demographic controls are introduced into the analysis, the suicide risk associated with these two groups becomes insignificant (neither significantly high nor significantly low). These two occupational groups are occupied predominantly by males. Earlier discussion (Chapter 2) suggested that, in the case of occupations which are occupied by demographic groups that are known to be at high risk of suicide (e.g. male), a true indicator of the suicide risk associated with that occupation can only be derived by introducing demographic controls into the analysis. The contrast between the findings of Stack (i.e. that, after demographic controls are introduced into the analysis, the suicide risk associated with these two categories of occupation is not statistically significant) and Kelly & Bunting (i.e. that these groups are associated with a significantly *low* level of suicide risk) lends credence to Stack's justification for incorporating demographic controls into analyses of occupation-related suicide risk.

In the case of carpenters and joiners, an occupational group which Kelly & Bunting's study indicates to be at significantly *high* risk (for the first time period), Stack also finds the associated suicide risk to be significantly *high*, but only on the basis of bivariate analysis. As in the case of the two occupational groups discussed in the previous paragraph, the risk is found to be *not* statistically significant after demographic controls are introduced into the analysis.

As discussion in this section has demonstrated, Stack's study and the two local studies reviewed are based on different theoretical premises and different methodologies. These differences give rise to findings that are, in many cases, discrepant. In the following section the implication of these discrepancies will be discussed, as well as their relevance to development of the research question for this thesis.

4. CONCLUSION

The two studies in England and Wales (Kelly et al, 1995; Kelly & Bunting, 1998) which are central to discussion in this chapter provide a ranking of the suicide risk associated with a range of occupations on the basis of the Provisional Mortality Rate for suicide computed for each occupation for the time period(s) relative to the study.

The model of analysis on which these studies is based does not incorporate controls for demographic factors apart from gender, although the authors explicitly acknowledge that other demographic factors such as age and marital status are predictor factors for suicide.

The American study (Stack, 2001) which is also central to discussion in this chapter is based on a logistic regression model of analysis of occupation-related

suicide risk. By introducing controls for four demographic covariates of occupation (i.e. gender, race, age and marital status), this model allows the risk of suicide associated with an occupation to be computed independent of these four demographic covariates of occupation.

Comparison of the findings from Stack's study with those of the two local studies is complicated by a series of factors. The first of these relates to the use of different measures for evaluation of suicide risk. The Provisional Mortality Ratio for suicide is used for the studies in England and Wales, while logistic regression odds ratios are used for the American study. The theoretical assumptions behind adoption of these indicators as a measure of suicide risk differ, and the indicators are not directly comparable with each other.

The second complication relates to the precariousness of suicide statistics over time, a phenomenon which was demonstrated with reference to the results of the earlier UK study (Kelly et al, 1995) and discussed in Chapter 2. This precariousness tends to confound the meaningfulness of attempts to compare occupation-related suicide statistics that are based on data from different time periods. This is one of the factors that complicates direct comparisons between the results of Stack's study, which is based on data for 1990, and the results of the England and Wales studies, which are based on data for 1982-1987 and 1988-1992 (Kelly et al, 1995) and 1982-1987 and 1991-1996 (Kelly & Bunting, 1998).

Comparison of results is also complicated by the use of different protocols for coding occupations. Definitions of occupational categories used in England and Wales studies are based on CO80 and SOC90, whereas definitions of occupational categories used in the American study are based on standard federal occupational codes. The use of different criteria for selecting occupational groups to be investigated is also likely to affect the comparability of results.

For the studies in England and Wales, suicides are defined as 'suicides and deaths from injury and poisoning undetermined whether accidentally or purposely inflicted' (Kelly & Bunting 1998: 29). But it is not clear whether 'undetermined deaths' are included in the definition of suicide used for Stack's study. The possibility that differing definitions of 'suicide' may have been used is another source of complexity in relation to the comparability of results.

Discussion in this chapter has suggested that a fall in the Provisional Mortality Rate for suicide for an occupation does not necessarily indicate a decrease in the rate of suicide for that occupation, but might indicate that the rate of deaths from other causes is increasing. This consideration is particularly relevant when comparing Provisional Mortality Ratios for Social Class I and II occupations with Provisional Mortality Ratios for Social Class IV and V occupations, since the PMR for suicide of an occupation may say more about the risk of *not* dying from other causes than about the risk of dying from suicide (Kelly & Bunting, 1992). Although findings for doctors, dentists and nurses have been shown in this chapter to be relatively consistent between one local study (Kelly & Bunting, 1998) and

Stack's study, findings for a series of lower socioeconomic status occupations are not consistent between the studies. As the discussion has shown, this discrepancy may to a considerable degree be related to the omission in local studies of controls for the demographic covariates of occupation (apart from gender). But it may also be related to the use of the Provisional Mortality Ratio as an indicator of suicide risk in local studies.

Discussion in this chapter has highlighted the complexities involved in assessing the reliability and validity of findings from investigations of occupation-related suicide by comparison with results from other studies. Discussion has also identified the complication that arises during comparisons when the studies use different indicators for measuring suicide risk. Discussion has highlighted the importance of controlling for the demographic covariates of occupation, and identified an area of theoretical interest where there appears to have been a shortfall in research in England and Wales, i.e. investigation of occupation-related suicide risk that is based on a multivariate model of analysis analogous to that employed by Stack.

Development of a model of analysis analogous to that used by Stack and the testing of it using local mortality data will be the focus for development of a research problem for this thesis. Such an attempt will depend fundamentally on locating and gaining access to a relevant source of primary data. The focus of the next chapter will be on an attempt to identify a source of individual-level mortality data suitable for such an analysis.

CHAPTER 4

AN OVERVIEW OF SOURCES OF MORTALITY DATA

1. INTRODUCTION

As explained in the final section of the previous chapter, the focus of this and the next chapter will be on the nature and sources of data suitable for use in developing a model of analysis of occupation-related suicide risk that is analogous to the model used by Stack (2001), i.e. a multivariate logistic regression model of analysis of suicide risk.

For the purpose of logistic regression analysis, the primary data must be *individual*-level (or 'unit record' level) mortality data, as opposed to *aggregate* data. The 'unit', for this purpose of the proposed analysis, is the person who died as a result of suicide. The logistic regression model of analysis used by Stack requires this individual-level data to include information about occupation, gender, marital status, race and age.

At the time I started this search, I was naïve about sources of mortality data in England and Wales and about legislation that regulates control of this type of data. But, as will emerge in this chapter, fairly early in the search I became aware that the search for mortality data was going to be complicated.

The authors of one of the studies discussed in Chapter 3 (Kelly & Bunting, 1998) are identified by the research report to be officers of the Office for National Statistics (ONS), so my search commenced with a request to ONS for information about the source of primary data on which this study was based. The reply advised that "there are limits on the use and disclosure of such data. You may get access to such data, but under certain conditions and this may take months" (F. Van Galen, ONS: personal communication, Dec. 2003). This was the first indication that the individual-level mortality data that was fundamental to an attempt to replicate Stack's methodology would not be easy - and perhaps not possible at all - to obtain.

Section 2 of this chapter offers a discussion of a number of organisations whose websites I explored during my search for mortality data. But, as will emerge later in this chapter and in Chapter 5, the search proved to be non-productive and valuable time was wasted in pursuing what eventually transpired to be a fruitless mission. However, this search revealed a series of insights into the inter-agency relationships that underpin the management of census-based and vital-events data

in England and Wales, a phenomenon that will be highlighted in the present chapter but discussed in greater detail in Chapter 6. The observation that access to individual-level mortality data is, in England and Wales, extremely difficult, began to explain the observation in Chapter 2 that empirical studies of suicide rates for multiple occupations are rare in England and Wales.

2. AN OVERVIEW OF SOURCES OF MORTALITY DATA

The discussion that follows is important not only from the point of view of demonstrating the process of my search for individual-level mortality data, but also from the perspective of illustrating the link between organisations and agencies that co-operate in the collection, processing and dissemination of census-based data and vital-events data.

2.1 UK DATA ARCHIVE (UKDA)

The UK Data Archive (www.UKDA[1]) is housed at the University of Essex and funded by the University of Essex, the Joint Information Systems Committee (JISC) of the Higher Education Funding Councils, and the Economic Social Research Council (ESRC). Its website claims it to be "a centre of expertise in data acquisition, preservation, dissemination and promotion" and "curator of the largest collection of digital data in the social sciences and humanities in England and Wales" (www.UKDA[1]). UKDA provides support for secondary use of quantitative and qualitative data in research, teaching and learning, and allows registered users to browse, analyse, or download data-sets from the website for most of the data series held by UKDA, or to order data that are not available to download.

At first sight, two of the datasets, neither of which was downloadable from the website, appeared to be of potential usefulness for my purposes: The *Historic Mortality Pack* (series number SN2902) (www.UKDA[2]), whose purpose is "to provide researchers with historical time series data on mortality", and the *Local Mortality Pack* (series number SN3625) (www.UKDA[3]) which "contains number of deaths by year, sex, age and ICD cause."

Both datasets have the Office for National Statistics as both "principal investigator" and "depositor" (www.UKDA[2]; www.UKDA[3]). At the time I investigated this website, the significance of this fact was not obvious to me, but will be discussed in greater detail in following sections.

Data for these files are held in age-groups. For most years the age groups are: under 1; 1-4; a series of 5-year age groups from 5 to 84; then all years 85 and over (www.UKDA[2]). This method of presenting data represents aggregation of individual-level data, and disguises the unit record level of data that is required for any analysis based on a logistic regression model. Even though I had impetuously

ordered these datasets, on examining their descriptions in greater detail, I realised that, because they do not contain individual-level data, they were of no usefulness for my purposes.

2.2 THE HUMAN MORTALITY DATABASE (HMD)

According to the website of HMD, the Human Mortality Database (www.HMD) was created to provide detailed mortality and population data to those interested in human longevity in order to facilitate research into its causes and consequences. The project contains detailed data for 20 countries, including England and Wales.

Although this project is based at the University of California, the fact that it contains data for England and Wales made it of potential interest. The raw data consist of death counts from vital statistics, plus census counts, birth counts, and population estimates from various sources. But, as in the case of the UKDA website, the database presents only aggregate, not individual-level data.

According to information given on the website, data for this project are provided by 'official sources' of the countries concerned. The website does not specify that, for England and Wales, the 'official' source is the Office for National Statistics (ONS). But later it my search for data it became apparent that this would indeed have been the identity of the 'official' source. As discussion in this and following chapters will demonstrate, the role of ONS is central to the management and dissemination of mortality data for England and Wales.

2.3 WORLD HEALTH ORGANIZATION (WHO) MORTALITY DATABASE

According to this website (www.WHO), the WHO Mortality Database contains data on registered deaths for individual Member States, including England and Wales, as "part of [WHO's] constitutional mandate to establish and maintain statistical services and provide information in the field of health". The data held on the WHO Mortality Database are, according to this website, "official national statistics" that are transmitted to the World Health Organization by the "competent authorities" of Member States. As discussion in Chapter 6 will demonstrate, the source of official national statistics for England and Wales is the Office for National Statistics. The fact that the source, on the WHO database, of statistics for England and Wales is the relevant "competent authority" is significant, once again, in highlighting the centrality of the role of the Office for National Statistics in holding and distributing official national statistics for England and Wales. This centrality, and the legislation that underlies it, will be discussed in greater detail in Chapter 6.

Data obtainable from the WHO Mortality Database are presented by age-group, sex, year and cause of death. The significance of the presentation by 'age-group' relates to the fact that, as in the case of the Human Mortality Database, access only to *aggregate* data is possible. From the point of view of an examination of the nature of mortality data available for England and Wales, the fact that access to individual-level data is precluded is significant. As discussion in this and subsequent chapters will demonstrate, legislation in England and Wales prevents access to individual-level mortality data in England and Wales, except by officers of the Office for National Statistics who are party to the Official Secrets Act. However, at the time that I was researching the websites described in the present chapter, I was naïve of this fact.

2.4 NATIONAL DIGITAL ARCHIVE OF DATASETS (NDAD)

The National Digital Archive of Datasets (www.NDAD) is operated by the University of London Computer Centre (ULCC) on behalf of National Archives (Public Record Office and Historic Manuscripts Commission). Its aim, according to its website, is to conserve and, where possible, provide access to a variety of computer datasets from Central Government departments and agencies.

NDAD holds two datasets, under the heading Historic Mortality Data Files (series CRDA/20)(www.NDAD[2]), that appeared, at first glance, to have potential value as a source of data for my own study. According to the website, the Historic Mortality Data Files database was originally created by the Office of Population Censuses and Surveys (OPCS) as a tool for researchers studying mortality in England and Wales. Two versions of this database are held by NDAD. The first version (series CRDA/20/DS/1) (www.NDAD[3]) contains annual data for England and Wales for the years 1901-1992 on the number of deaths by cause of death, sex, and age group, together with annual populations estimates by sex, year and age-group. The second version (series RDA/20/DS/2) (www.NDAD[4]) covers the period 1901-1995, and replaces the earlier version. According to the NDAD website, the second version of the Historic Mortality Data Files was redesigned by ONS and issued (by ONS) to the public on CD-ROM under the title "Twentieth Century Mortality Files" in 1997 (www.NDAD[5]; www.ONS[1]). After 1997, CD-ROMs were issued annually to cover the years from 1996 until 1999 and in 2003 a revision of the Twentieth Century Mortality Files was issued on CD-ROM under the title 20th Century Mortality (England & Wales, 1901 - 2000)(www.ONS[1]). This database covers the period 1901 to 2000 and incorporates revised population estimates based on the 2001 Census.

Development of the Historic Mortality Files was inspired by an initiative by the World Health Organization in the mid 1970s to create a database of death rates from 1950 onwards for a number of countries by year, sex and a limited range of

causes. In 1979 the Office of Population, Censuses and Surveys (OPCS) began to develop a mortality database for England and Wales which was designed to facilitate studies of mortality which require national rates for comparative purposes. The database was constructed from "readily available sources in published form and on computer" (www.NDAD[5]). One of these sources is information collected from death registration data. This information is processed in accordance with mandates legislated by the Populations (Statistics) Act, although I was not aware of this fact, nor of its relevance to my attempts to access individual-level mortality data, at the time I investigated the NDAD website. In terms of this legislation, data which have been derived from birth and death registration can never be released on an individual basis to parties external to the Office for National Statistics. A more extensive discussion of this legislation and its relevance to research on suicide will be offered in Chapter 6.

Discussion of NDAD highlights, once again, the link between agencies which hold mortality data and the apparent sovereignty of the Office for National Statistics over individual-level mortality data (and other categories of data) in England and Wales – although the significance of this observation was not apparent at the time I first explored this website. This link is fundamental to understanding the control in England and Wales of mortality data derived from death certificates and the legislative basis for this phenomenon, and will be discussed in greater detail in later chapters. But, for the purposes of discussion in the present chapter, it is relevant to note, firstly, that the Historic Mortality Data Files (www.NDAD[2]) and the Twentieth Century Mortality Files (www.ONS[1]) do not contain information on occupation and, secondly, present only aggregate data on mortality. They are thus, on two counts, not suitable sources of data for multivariate logistic regression analysis of occupational suicide risk.

2.5 CENTRE FOR LONGITUDINAL STUDY USER AND INFORMATION SUPPORT (CeLSIUS)

Discussion in this section further demonstrates the link between agencies that hold mortality data for England and Wales, in this case the link between the Office for National Statistics and the Centre for Longitudinal Study Use and Information Support (CeLSIUS), which is funded by the Economic and Social Research Council (ESRC).

The Longitudinal Study (LS) (www.ONS[2]) was designed by the Office for National Statistics as a continuous study based on a random 1 per cent sample of the population of England and Wales, using four specified annual birthdates as the selection criterion, and containing linked census and vital events data (births, deaths, widow(er)hoods, migration, enlistments, and entries to long-stay hospitals) on a 1% sample of the population (Hattersley & Creeser, 1995).

The study constitutes a complete set of individual-level census records, linked between successive censuses (Hattersley & Creeser, 1995). It was started in 1974 with a sample drawn from the resident population of England and Wales as enumerated in the 1971 census. Four equidistant dates in the year were used to select an approximate 1% sample of the population that could be followed over time. The four dates were used to update the sample at the 1981 and 1991 censuses and in routine event registrations.

Selection onto the Longitudinal Study is by birth date (Hattersley & Creeser, 1995). New LS members enter the study through birth and immigration, and existing members leave through death and emigration. Further members from the 2001 Census are due to be added to the Longitudinal Study in September 2004. The Longitudinal Study thus represents a continuous sample of the population of England and Wales, rather than a sample taken at one point in time. At present it includes records for approximately 1 million sample members (www.ONS[2]). Data about Longitudinal Study members are derived from responses to census questions and from data routinely collected by the Office for Population Censuses and Statistics (OPCS).

The Longitudinal Study is held by the Office for National Statistics (ONS), but the Centre for Longitudinal Study Information and User Support (CeLSIUS) serves as the information and support team for researchers who wish to access data held on the Longitudinal Study (www.CeLSIUS[1]). ONS's policy is that the Longitudinal Study should be used "as widely and easily as is consistent with maintaining the confidentiality of the data" (www.CeLSIUS[2]). The legislation underlying ONS's commitment to 'confidentiality' and the full implications of this commitment were unknown to me at the time I considered the Longitudinal Study as a potential source of primary data. Only later, in the course of analysis of the legislation underlying ONS's policy on 'confidentiality', did the reasons for and full implications of this policy become apparent, as will be discussed in Chapter 6. But, for the purpose of discussion in the present chapter, it is only relevant to note that, on the basis of statutory vetos, access to individual-level data can never be granted to parties external to ONS.

The CeLSIUS website offers guidelines for applying for access to the Longitudinal Study, but specifies that only tabulation statistical summaries or *aggregated* data, that can be analysed or modeled by the user, can be provided to users (www.CeLSIUS[2]). The website specifies that analysis of *individual* records can be carried out *only* on the ONS computer and *only* by CeLSIUS staff, and only statistical abstracts of these analyses can be issued to users (www.CeLSIUS[2]).

Optimism generated by finding a website that offers access to a database that includes individual-level data had waned during my examination of the CeLSIUS website and discussions with ONS and CeLSIUS staff. The purpose of the search for mortality data had been to attempt to replicate Stack's methodology using local data and thus to address an (apparently) neglected area of research in occupation-

related suicide. Performing such an analysis requires, as its starting point, access to *individual-level* mortality data, and such data must, necessarily, include occupational data. It had begun to seem that gaining access to such data would, at the least, be difficult, but might be impossible.

However, a conversation between one of my academic advisors and a member of CeLSIUS staff led to the impression that the conflict between my requirement to base the analysis on individual-level data and ONS's policy not to permit direct access to this data could potentially be resolved by CeLSIUS's facilitating indirect access to the data by conducting the statistical analyses on their own computers in terms of specifications that I would provide. It was on the basis of this possibility that I decided to submit a proposal to CeLSIUS for access to mortality data held on the Longitudinal Study. The proposal and its outcome are the subject of Chapter 5.

In the next section of this chapter, a brief review is offered of a research option that emerged as a possible alternative to the proposal based on logistic regression analysis of mortality data.

3. A REVIEW OF THE RESEARCH OPTIONS

Stack's study on occupation-related suicide, as discussed in Chapter 3, was based on access to mortality data for the 21 states of the United States that report occupational data in the 1990 National Mortality Detail File (Stack, 2001). The Longitudinal Study, which is the only potential source of mortality data for a similar study in England and Wales, represents only 1% of the population of England and Wales. Although the size of this sample is not comparable to the population available to Stack, the purpose of my own analysis would be to attempt to replicate Stack's *methodology* rather than an attempt to replicate every feature of the study. An attempt to compare results of my own analysis with those of Stack's would be confounded in any case by a variety of factors (as discussed in Chapter 3) *othe*r than population size, such as the use of differing occupational classification systems.

But, even if my application for access to data from the Longitudinal Study were approved (which was unlikely), I envisaged a number of other complications in managing and categorizing the data for use in my own study. A fundamental consideration in use of a longitudinal data source is the effect of changes in the coding of events over time. For example, the International Classification of Diseases (as discussed Chapter 3, section 5.1) has been revised twice during the history of the Longitudinal Study – although, as demonstrated in Chapter 3, the transitions between ICD-8 and ICD-10 have not had a significant effect on classification of suicide deaths.

However, in the case of revisions of the Classification of Occupations, the effect of transitions cannot be ignored. Occupations are classified according to the

Classification of Occupations, which has also been revised twice during the time spanned by the Longitudinal Study, i.e. since the early 1970s (see Chapter 2 and Table 13). These changes must be accounted for by recoding the variables to a common base (www.CeLSIUS[4]). This process requires knowledge of and expertise in the use of occupational coding frames and is central to managing and categorising occupational data from the Longitudinal Study. Thus, even in the unlikely event that my application to CeLSIUS were successful, I anticipated that my limited proficiency in the use of these classification systems was likely to be a fundamental drawback.

Table 13 Occupational classifications used in England and Wales and year of use

	Code	Title
1970	CO70	Classification of occupations
1980	CO80	Classification of occupations
1990	SCO90	Standard classification of occupations
2000	SOC2000	Standard occupational classification

In view of this potential difficulty, I now drew up a revised version of the earlier objective (i.e. an attempt to replicate Stack's methodology for analysing occupation-related suicide risk, using local data). The goal would be to develop and test a multiple regression model of risk for suicide for only *one* occupation. In light of discussion in Chapter 3, in which doctors have consistently been shown by many research studies to be at high risk of suicide, this would be the occupation investigated in the modified version of the study. This choice had the further merit that this occupation has been less affected by revised versions of occupational classification systems over the history of the Longitudinal Study than have definitions of occupations that have either become redundant during this time (e.g. comptometer operators) or have emerged and evolved during this time (e.g. occupations relating to information technology).

Discussing the viability of this option with one of my research supervisors prompted me to review my original research interest (as detailed in Chapter 1): "What are the characteristics of macrosocial influences that negatively impact on individual psychological wellbeing?". In light of the probability that application to CeLSIUS for access to individual-level data, even for only 'doctors', would be unsuccessful, I reflected that it would be wise to revisit my original research interest and reformulate my research objective. The option that emerged is discussed below. However, since I intended, on the advice of one of my research advisors, to proceed with the application to CeLSIUS anyway, I shall review this option only briefly.

Stack's *multicausal heuristic model of occupation-related suicide,* which has been discussed in earlier chapters, identifies four aspects of occupation that may potentially influence the suicide risk of any one occupation: stress internal to the

nature of the occupation itself, pre-existing psychiatric morbidity, opportunity factors, and demographic factors.

My earlier research objective (i.e. an attempt to replicate Stack's multivariate regression model for investigating occupation-related risk, using local data) centres on the demographic dimension of the full model. But discussion in the present chapter has highlighted the difficulty of obtaining relevant local data. The revised research objective shifts the focus to the first dimension of this model (i.e. stress internal to the nature of the occupation itself) and is oriented towards my original research interest (i.e. what are the characteristics of macrosocial influences that negatively impact on individual psychological wellbeing?). For the purposes of the revised research objective, this question would centre round a relationship between the following variables:

- The level of 'control' associated with an occupation (which would represent 'macrosocial influences') and
- The suicide risk associated with that occupation (which would be employed as an indicator of 'individual psychological wellbeing').

The aim of the study would be to develop an instrument for measuring variations in 'occupational control'.

Perfunctory reflection suggests the following indicators (inter alia) of 'control' associated with any particular occupation:

- Central government directives
- Government inspections
- Government directed administration
- Statutorily mandated membership of professional bodies
- Policies and legislation concerning working hours and remuneration.

Deriving a scale measuring 'occupational control' would enable the suicide risk of that occupation to be computed in terms of the relationship between the level of control associated with any particular occupation and the rate of suicide associated with that occupation.

As discussion on the use of the Proportional Mortality Ratio as an indicator of suicide risk has demonstrated in earlier chapters, the higher rate of suicide that tends to be associated with Social Class I and II occupations may be a more significant indicator of a low rate of death from 'other causes' than of a high rate of death from suicide. This observation highlights the significance of controlling for socio-economic status in analyses of occupation-related suicide.

The review of theoretical literature in Chapter 2 demonstrated that medical practitioners have consistently been found to be at high risk of suicide, while the

deficit of literature relating to suicide amongst legal practitioners (another Social Class 1 occupation) suggests that the suicide risk associated with this occupation is neither significantly high nor significantly low. Medical practice and legal practice thus represent socially equivalent occupations, each associated with a different category of suicide risk. Interviewing a number of practitioners from each of these occupations would yield information about the comparative conditions under which these occupations practice, with socioeconomic status having been controlled for. This would enable construction of 10 questions for the purpose of deriving a 20 point scale of occupational control, which could then be tested on other occupations.

There is a tendency in the secondary literature towards arguing that ease of access to means of committing suicide as well as expertise in the use of those means is a key factor in making an occupational group vulnerable to suicide (e.g. Kelly & Bunting, 1998). The presence of doctors, nurses, dentists, pharmacists and chemical scientists amongst high-risk occupational groups appears to lend credence to this argument. However, my own hunch is that the level of control (or lack of) that a practitioner has over the conditions under which the occupation is practised is as significant a contributory factor to suicide risk as is ease of access to means of committing suicide. Developing a scale for measuring 'occupational control' would enable this hunch to be tested by drawing up a correlational graph to demonstrate the relationship between the relative suicide risks of a range of occupations against the levels of 'occupational control' associated with those occupations.

4. CONCLUSION

Stack's multivariate logistic regression model for investigating occupation-related suicide risk is based on individual-level mortality data, and an attempt to test this model using local data depends, fundamentally, on access to individual-level mortality data. As this chapter has demonstrated, access to such data for England and Wales appears not to be possible, except for authorised officers of the Office of National Statistics (ONS). This chapter has also demonstrated a link between agencies that hold mortality data for England and Wales, both locally and abroad, as well as the apparent supremacy of ONS in controlling access to this data. The legislative basis for this supremacy, the Populations (Statistics) Act and the Census Act, will be discussed in greater detail in Chapter 6.

Despite the legislative veto against ONS sharing data from the Longitudinal Study with outside parties, a conversation between one of my academic advisors and an official of ONS led to the impression that my application for access to individual-level mortality data could potentially be approved on the basis of staff at ONS performing the logistic regression analyses in terms of specifications provided

by myself. It was on the basis of this possibility that I decided to proceed with my application to CeLSIUS, as will be discussed in the following chapter.

In view of the possibility that my application would *not* be approved, a possible alternative research question for my thesis has been described briefly in the present chapter. This question centres round the following dimensions:

i) Exploring occupational-control issues with representatives of two socially-equivalent professions in which there are contrasting suicide rates, and
ii) Collecting and developing items for incorporation into a reliable and valid scale for measuring the level of 'occupational control' of an occupation.

A scale of this nature would enable the level of 'occupational control' for individual occupations to be computed. The use of such a scale would allow the drawing up of a correlational graph to demonstrate the relationship between the level of occupational control associated with a range of occupations and the suicide risk associated with those occupations. For this computation, the 'occupational control' scores would need to be computed for a range of occupations. I did not envisage, at the time of drafting this idea, that this would form part of my own study but rather that my study would provide the basis for future investigations.

During discussion about the two research options outlined in Section 3, one of my academic advisors advised against the second option, expressing the view that such a project was too extensive for a master's level thesis. At his suggestion, I decided to proceed with an application to CeLSIUS, despite the evidence that it was unlikely to be successful. The process and the outcomes of my application are described in the following chapter.

CHAPTER 5

EVALUATION OF THE LONGITUDINAL STUDY AS A VENUE FOR OCCUPATION-RELATED MORTALITY DATA

1. INTRODUCTION

In deciding to proceed with an application to CeLSIUS, I had to consider the possibility that the outcome might not be favourable. I was only to a degree comforted by my research supervisor's reassurance that the focus in evaluating a master's level project is on the assessment of theoretical process and reflexivity demonstrated in the thesis, as much as on empirical outcomes. For, if my application to CeLSIUS were *not* successful, there was the possibility that there might be no empirical outcome at all.

But, even considering the possibility that my application might unexpectedly turn out to be successful, I still had a number of reservations about the feasibility of an attempt to develop and test a statistical model analogous to Stack's (as described in Chapter 2). These reservations related to a large extent to the complexities involved in recoding variables such as 'occupation' and 'suicide' to a common base (a complication which was discussed in Chapter 4).

The draft application was framed in terms of a request for indirect access to Longitudinal Study data, at an individual-level, with staff at CeLSIUS running the logistic regression analyses on the basis of specifications that I would provide. This process would require ongoing collaboration between staff at CeLSIUS and myself in order to secure meaningful results. The application (see Appendix 1) framed the questions I would be seeking to investigate as follows:

i) Are there occupational groups in England and Wales that have a significantly high or significantly low risk of suicide in relation to comparable groups in the general population?
ii) How do the demographic variables of age, sex, marital status and race mediate the risk of occupation-related suicide?
iii) Does the process of introducing controls for the demographic covariates of occupation support or refute the findings of Kelly & Bunting's (1998) study?

Chapter 5 centres around the application to CeLSIUS and the ensuing discussions with CeLSIUS staff as an attempt was made to reach a compromise to the conflict between my requirement to base the analysis on data at individual-level, and the CeLSIUS policy to not permit direct access to such data.

As this chapter will demonstrate, the outcome was not favourable and the proposal ultimately proved to be unviable. However, much time had elapsed since my first contact with ONS and CeLSIUS, and this left me in the precarious position of not being left with sufficient time to backtrack to the alternative research option detailed in Chapter 4, section 3 (i.e. to develop a scale to measure 'occupational control').

The present chapter offers an overview of the discussions I had with CeLSIUS staff and a discussion of the rationale for deciding not to take the application further.

2. DISCUSSIONS WITH CeLSIUS

My application was submitted in February 2004 (see Appendix). The application acknowledged that, in consideration of CeLSIUS's commitment to 'confidentiality', *direct* access to individual-level data would not be expected, but *indirect* access to this data was requested. This request was based on an earlier discussion between a CeLSIUS officer and one of my academic advisors, in which the impression had been given that CeLSIUS might consider facilitating *indirect* access to individual data by running the statistical analyses according to specifications I would provide. The viability of the proposed study depended fundamentally on CeLSIUS's response to this request.

The initial response from CeLSIUS did not address this request directly. Discussion in the e-mail centred on suggestions for modifying the proposal in such a way as to allow the analysis to be based on *aggregate* data (J. Buxton, CeLSIUS: personal communication, Feb, 2004), and gave the impression that, despite the impression given earlier, CeLSIUS was not entertaining the request for indirect access to individual-level data.

Without clarification about this fundamental consideration (i.e. access to individual-level data), the other issues raised in the e-mail were of secondary importance. In response to this e-mail (personal communication, March 2003), I re-iterated that the analysis described in the proposal was based *fundamentally* on individual-level data and that "I understood, from a phone conversation you had with one of my academic advisors early in February, that you had suggested a possible way around *my* requirement to base the analysis on individual-level data and *your* policy not to permit direct access to this data, and incorporated your suggestion in my application as follows: 'CeLSIUS's contribution would be to perform the analyses, on the basis of my specifications.' "

CeLSIUS's response confirmed that CeLSIUS "would not be able to provide detailed occupational data at the level you discuss" and suggested that "If you are able to work with aggregated data, then we may be able to move forward" (J. Buxton, CeLSIUS: personal communication, Feb, 2004). A later e-mail re-iterated that "we cannot give you a dataset for you to analyse that includes information at the level of data you asked for" on the grounds that "there are strict confidentiality requirements that need to be met" (E. Grundy, CeLSIUS: personal communication, March 2004).

Shortly afterwards, I was advised that the Longitudinal Study Research Board was to meet within the next few days and that my proposal had not been submitted for consideration as "it was clear that the request as specified there would not be granted" (E. Grundy, CeLSIUS: personal communication, March 2004). This e-mail suggested, however, that my request could still be tabled at the meeting "provided the specification is settled very soon". This e-mail re-iterated that "a rare cause of death for specific occupational groups is not going to be permissible", but suggested that "broader groupings, such as medical practitioners, nurses and dentists together, could be considered". Medical practitioners, nurses and dentists are categories of occupation that have each consistently been identified as being at high risk of suicide (see Chapter 2). Collapsing these occupations into one group would negate the potential for comparing the risk of suicide associated with each of these occupations separately and, for this reason, the suggestion offered by CeLSIUS was not viable for the purposes of addressing the research questions on which my proposal to CeLSIUS was based (as detailed in section 1 of this chapter).

In meeting with my research supervisor to discuss the implications of this last e-mail, we agreed that, even if there *had* been time to arrange for a request to be tabled at the Longitudinal Study Research Board meeting scheduled for a few days hence, it was unlikely that my specifications would be approved at that meeting or at any time in the future.

This realisation served as the cue for the discussion that follows (in the remainder of the present chapter and as the focus of Chapter 6) on the statutory factors that underlie the differences between local and American practices relating to the collection, processing, holding and sharing of individual-level mortality data (as well as other vital-events data and census data). The remainder of the present chapter centres on CeLSIUS policy in relation to privacy of Longitudinal Study subjects, and on a preliminary discussion of related policies in the United States.

3. CeLSIUS POLICY ON CONFIDENTIALITY AND ITS RELEVANCE TO STUDIES OF OCCUPATION-RELATED SUICIDE

According to the CeLSIUS website, "Protecting confidentiality is of paramount importance when releasing LS outputs. No output may be released that

might allow a user to identify an individual within the dataset ... Therefore there are various constraints governing what can and cannot be released, and the user must bear these in mind when making their request" (www.CeLSIUS[5]).

The veto against access to individual-level data and the practice of releasing only aggregated data is one of the practices that protects against potential disclosure of identity of subjects included in the Longitudinal Study. Another practice relates to cell counts in tables. According to the CeLSIUS website (www.CeLSIUS[5]), "A cross-tabulation might have a potential for disclosure if certain cells contain very low numbers of individuals, and if the combination of variable categories represented by that cell are recognisable characteristics that might be used to identify an individual (or individuals)."

In the case of data requests that CeLSIUS considers to be "potentially disclosive", the cross-tabulation may not be published until some form of aggregation is done to increase the cell counts. This is usually done by recoding one or more of the variables so that it has fewer categories (www.CeLSIUS[6]; www.CeLSIUS[7]). The characteristic of being "potentially disclosive" is attributed to "certain combinations of variables", in particular to combinations of variables relating to ethnicity, minor occupational groups or sub-regional geography (www.CeLSIUS[6]; www.CeLSIUS[7]). This is because there are likely to be low counts in the rarer categories of these variables.

A table is considered to be *not* "potentially disclosive" if no cell count is less than 3 (www.CeLSIUS[6]). However, in the case of suicides for single occupations, the policy is that the cell count must not be less than 5 (F. Van Galen, ONS: personal communication, June 2004). Since the Longitudinal Study is based on a 1% sample of the general population, a cell count of 5 corresponds to roughly 500 in the general population. This means, in the case of studies of occupation-related suicide, that, even if the model of analysis is designed to accommodate aggregate data (which Stack's logistic regression model is *not*), an occupation must be characterized by at least 500 suicides since 1971 (i.e. since the start of the Longitudinal Study) before even *aggregate* data on that occupation can be released to an external user.

For the age group 16-64 (the age group of interest in studies of occupation-related mortality), the Longitudinal Study represents approximately 46 suicide deaths in any one year (A. Ross, CeLSIUS: personal communication, Feb. 2004). This means that the 30 year time-span of the Longitudinal Study represents fewer than 1500 suicide deaths of interest to occupation-related studies. Thus, the suicide count for any *one* occupation is unlikely to be as high as 500, the critical number that must be reached before even *aggregate* data on that occupation can be published to an external user.

The discussion in this paragraph gives rise to the observation that, even for analyses of occupation-related suicide based on *aggregate* data, the Longitudinal Study is unlikely to be an adequate venue for data.

4. CONCLUSION

As discussion in this chapter has demonstrated, access to individual-level mortality data held in the Longitudinal Study is, in the case of users external to ONS and CeLSIUS, vetoed on the grounds of 'confidentiality'. This has fundamental implications for the viability of the project proposal which has been the focus of the first five chapters of this thesis, a logistic regression model of analysis of occupation-related suicide risk which is based, fundamentally, on individual-level primary data.

Confidentiality of personal data in the Longitudinal Study, the central source of census and vital-events data for England and Wales, is protected by the Census Act and by the Population (Statistics) Act. These statutes place constraints on the way data are held as well as the way in which they can be extracted (Hattersley & Creeser, 1995), and give rise to the veto against the release of individual-level mortality data held on the Longitudinal Study to the public.

The finding that it would not be possible to obtain the individual-level mortality data which is central to performing the statistical analysis which is the focus of this thesis, served as the cue for a decision to explore the legislative mechanisms that underlie policies in United States and England and Wales relating to sharing of mortality data with the public.

By contrast to the situation in England and Wales, detailed information on mortality is freely accessible in the United States. The example demonstrated by the United States illustrates that it is not beyond the realms of reality for mechanisms to be implemented that protect the rights of subjects to confidentiality, while at the same time enabling access to national data by researchers.

The data for Stack's study (as discussed in Chapter 3, section 2) were drawn from the National Mortality Detail Files (www.ICPSR[2]), which were developed by the National Center for Health Statistics (NHCS), which is the United States' principal health-statistics agency (www.NCHS[1]). As in the case of ONS's Longitudinal Study (which is derived from census and vital-events data), NCHS's data-holdings were originally collected for specific research or administrative purposes. But, by contrast to the case of the Longitudinal Study, NCHS's data-holdings are recognised to have research potential that outlives the original purposes, and NCHS's authorizing legislation mandates that data be made publicly available for secondary analysis (www.HHS[2]). The different legislative mechanisms that underlie the different policies on public access to national data will be discussed in greater detail in Chapter 6.

As this chapter has demonstrated, it is not possible to access individual-level mortality data through the Longitudinal Study. But, just as significantly, access to

aggregate data is also limited. An occupation must be characterized by at least 500 suicides before even aggregate data on that occupation can be released to an external user. But, since the Longitudinal Study is represented by less than 1500 suicide deaths in the 16-64 age-group, it is unlikely that many (if any) occupations would be associated with numbers that would allow the release of data on that occupation. This limitation places considerable restrictions on the potential of the Longitudinal Study as a venue for *aggregate* data for occupation-related suicide studies, quite apart from being futile as a venue for *individual*-level data.

A document published by the Office of Population Censuses and Surveys, which is available on the ONS website (Hattersley & Creeser, 1995) serves as an overview of the Longitudinal Study and as a "comprehensive technical reference document" (Hattersley & Creeser, 1995). This document discusses, inter alia, the Longitudinal Study's data sources, its design and methodology, and the legislation which regulates it. It seems a pity that, in emphasising the central commitments to "confidentiality", neither the CeLSIUS website nor the ONS website direct the reader immediately and authoritatively to this document. It is also a pity that this document wasn't referred to in correspondence with CeLSIUS referred to in Section 2 of this chapter. It was only after extensive research for information to explain the policy of "confidentiality" relating to the Longitudinal Study that I came across this document and discovered that my research proposal, by virtue of being formulated in terms of access to *individual*-level mortality data, had been doomed from the start.

A discussion of relevant legislation and its implications for access to individual-level mortality data and studies of occupation-related suicide follows in Chapter 6.

CHAPTER 6

MORTALITY DATA AND LEGISLATION

1. INTRODUCTION

As Chapter 5 demonstrated, the Longitudinal Study appears to be, at one and the same time, the primary venue of mortality data for England and Wales for researchers, a futile venue for *individual*-level mortality data, and (in the case of studies of occupation-related suicide) a fundamentally limited venue for *aggregate* mortality data.

The possibility that there might be legislative impediments to testing a logistic regression model of analysis of occupation-related suicide using local data did not cross my mind at the time I decided to base my project proposal on an attempt to test Stack's multivariate model of analysis (which is discussed in Chapter 2) using local data. As discussion in earlier chapters has demonstrated, by the time I submitted the proposal (as discussed in Chapter 4) to CeLSIUS for access to primary data held on the Longitudinal Study, I was aware of ONS's and CeLSIUS's legal obligation to protect the identity of Longitudinal Study members, but naïve of the complexities that underpinned this obligation. I recognized that the outcome of my application might not be successful, although l held out the hope that CeLSIUS might facilitate *indirect* access to the relevant data (in the manner described in Chapter 5, section 2).

In responding to my proposal, CeLSIUS staff referred to the issue of 'confidentiality' and the statutory basis for it, i.e. the Census Act and the Population (Statistics) Act. But I believe that a definitive and authoritative pointer to the ONS publication *Longitudinal Study 1971-1991: History, organisation and quality of data* (Hattersley & Creeser, 1995) and, in particular, to the section entitled "Confidentiality and access" would have been helpful. This document outlines the statutory constraints on release of individual-level data from the Longitudinal Study and specifies that primary analysis of data can only be conducted within the confines of the Office for National Statistics and then only by persons who have signed the Official Secrets Act. Reference to this document would have established, early in the proceedings, that my application for access to individual-level data was fundamentally a 'mission impossible' - unless CeLSIUS enabled indirect access to the data, as there had been some indication they would. By the time I came upon the document referred to in the preceding chapter and realised

that my research proposal, as it stood, had been fundamentally unviable from the beginning, too much time had elapsed to allow me to backtrack to the alternative research proposal outlined in Chapter 4, section 3.

Chapter 6 seeks to provide an overview of the legislative complexities that underlie protection of personal data and issues relating to sharing of data. It will also examine differences in legislation that make secondary analysis of census-based data and vital-events data possible in the United States for the public but unachievable (for the public) in England and Wales. It will then go on to offer a more general evaluation of the project and the research decisions I made that resulted in the failure to produce the anticipated empirical outcomes, as well as recommendations for future research.

The chapter commences with a discussion of the legislative basis for restrictions on access to data held in the Longitudinal Study. The first section demonstrates that, in the United States, there are no similar restrictions on access by the public to census-based and vital-events data, and reviews the legislative factors that underlie this difference. The purpose of this discussion is to highlight relevant commonalities and differences between the United States and England and Wales in relation to legislation and philosophies concerning sharing of national data with the public, and especially how this relates to access to individual-level mortality data.

The final section of this chapter offers an overall review of this thesis and highlights some areas of research which warrant further investigation.

2. RESTRICTIONS ON ACCESS TO LONGITUDINAL STUDY DATA

Information held on the Longitudinal Study is protected by two statutes: the Census Act and the Population (Statistics) Act. The relevance of the statute depends on the source of the data. Longitudinal Study data which are derived from census sources are protected by the Census Act of 1920, amended by the Census (Confidentiality) Act 1991. This statute lays out the requirement that personal data collected from census respondents not be disclosed to any person except Census Office staff involved in the initial processing of the census (Rees, Martin & Williamson, 2002).

The Act further stipulates that, although names and addresses will be recorded on the census form, they will remain there unused for 100 years (Hattersley & Creeser, 1995). Software is used to lift data from the census form and process it electronically but, apart from the unit postcode part of the address and the collection area code, the name and address of the respondent are excluded. Census data are thus anonymised at the point of capture onto the Census Office database.

The Census Office uses a further two strategies to protect census data from the risk of disclosure. The first involves both physical security of the Census Office computers and the policy of not connecting the Census Database to computer networks. The second involves systematic vetting of outputs (such as tables and datasets for publication) for unique and sensitive classifications that could potentially result in subjects being identified (Rees, Martin & Williamson, 2002). Although direct access to census forms is vetoed for 100 years, the Census Act does allow for certain datasets, after the data have been subject to the protection devices of vetting, sampling, broad coding, and licence agreement, to be released to authorized users, known as 'census data resources' (Rees, Martin & Williamson, 2002). The Longitudinal Study is one such 'census data resource' and the respective 'authorised' (i.e. 'licensed') user is the Office for National Statistics.

Data on the Longitudinal Study which have been derived from birth and death registration are protected by the Population (Statistics) Act (1938). This Act regulates access to the confidential particulars supplied at birth- and death-registration and stipulates that these data can be used for statistical purposes only and can never be released on an individual basis to external parties. Birth and death and other vital-events data (i.e. widow(er)hoods, cancer registrations, migration, enlistments, entries to long-stay hospitals) held on the Longitudinal Study have not been anonymised and "this explains all the security around the Longitudinal Study unit and the constraints on access to data" (F. Van Galen, ONS: personal communication, June 2004).

In summary, census data and vital-events data (including mortality data) can never be released to the public at individual-record level (even though census data are held in the Longitudinal Study in anonymised form). Data can be released to outside users only in aggregated form, i.e. in the form of statistics or tabulations (Hattersley & Creeser 1995). These data can be subject to secondary analysis only within the confines of the Office for National Statistics, and then only by persons who have signed the Official Secrets Act.

It is the practices described above, and the statutory basis for them, that gives rise to the situation whereby access to individual-level census and vital-events data (such as mortality data) by the public is fundamentally curtailed in England and Wales. This is by stark contrast to the situation in the United States, where accessibility to such data is not only possible, but mandated.

The following section provides a review of the legislation and practices in the United States that facilitate access to individual-level census and vital-events data, and which enabled Stack to conduct studies of occupation-related suicide on the basis of relevant individual-level mortality data (as discussed in earlier chapters).

3. ACCESS TO INDIVIDUAL-LEVEL POPULATION DATA IN THE UNITED STATES

The primary data on which Stack's study on occupation-related suicide (see Chapter 3, section 2) was based were drawn from the National Mortality Detail Files (www.ICPSR[2]), which were developed by the National Center for Health Statistics (NCHS) (www.NCHS[1]) and which are held by the Inter-University Consortium for Political and Social Research (ICPSR) (www.ICPSR1).

The National Centre for Health Statistics (NCHS) was created in 1960 out of a merger between the National Health Survey and the National Office of Vital Statistics (www.HHS[2]). It is the United States' principal health-statistics agency and is an organisational part of the Centers for Disease Control (CDC). CDC, in turn, falls under the Department for Health and Human Services (www.HHS), and is the lead federal agency for (inter alia) developing and applying diseased prevention and control (www.CDC[1]).

One of NCHS's primary objectives is to guide public health and policy, through the collection, analysis and dissemination of data on all aspects of the health of the United States population (www.NCHS[4]). NCHS has two major types of data systems: systems based on populations, containing data collected through personal interviews or examinations; and systems based on records, containing data collected from vital and medical records (www.NCHS[1]). More than 70 data files, including the National Mortality Detail Files, have been deposited by NCHS with the Inter-university Consortium for Political and Social Research (ICPSR) for public use.

ICPSR, which was established in 1962, is a not-for-profit organisation, with a membership basis of more than 500 colleges and universities, both in the United States and abroad. It maintains and provides access to an extensive archive of social-science data for research. The ICPSR data holdings consist mainly of raw data derived from surveys, censuses and administrative records. In common with NCHS, ICPSR is committed to the goal of "equitable access to national data" for research and teaching purposes (www.CDC[3]; www.ICPSR[1]), and these files can be readily downloaded from the website for secondary analysis. ICPSR's holdings of NCHS data-files include (inter alia) data from the National Health Interview Survey, National Ambulatory Medical Care Survey, National Hospital Discharge Survey, and the National Health and Nutrition Examination Survey as well as the National Mortality Detail Files (www.ICPSR[1]).

Although NCHS's data-holdings, such as the National Hospital Discharge Survey, were originally collected for specific research or administrative purposes, they are recognised to have research value that outlives the original purposes, and NCHS's authorizing legislation mandates that data held by NCHS be made publicly available for secondary analysis (www.HHS[2]). This is by stark contrast to the situation (as discussed in Chapter 5) whereby the Census Act and Population

(Statistics) Act legislate against the sharing of census and vital-events data with the public, even after the data has (in the case of census data) been anonymised.

The same legislation that requires NCHS to make its data-holdings available to the public also requires NHCS to safeguard the identity of individuals or establishments represented in its data systems (www.NCHS[4]). But the right of data-subjects to 'privacy' has to be weighed against the right of the public 'to know' (www.NCHS[4]). To this end, files for publication undergo a rigorous confidentiality review by the NCHS Confidentiality Officer and the Disclosure Review Board and are not released until the micro-data have been edited and judged to be free of potentially identifiable information (www.NCHS[4]; www.HHS[2]). Data items that could potentially be used to identify individual respondents are removed, masked, or collapsed in the public-use versions of the datasets that are released to the research community.

As a further measure for protecting the privacy of subjects, NCHS stipulates (in terms of the Public Health Service Act)(www.PHSA) that data be used by the public only for the purposes of statistical analysis and reporting. The data may not be used for the purpose of learning the identity of any person or establishment. Analysts are required, if identities are inadvertently discovered, to notify the director of NCHS (www.ICPSR[1]). But, in any case, the risk of identities being discovered has been further protected against, since 1989, by the omission from files of the date of birth or date of death of the decedent, as well as restriction of geographical details (www.ICPSR1).

The Mortality Detail Files (www.ICPSR2), from which Stack drew primary data for his study, is a data collection which describes every death registered each year in the United States from 1968 to 1991. Data are drawn from the vital statistics offices of each state as well as the District of Columbia for the years 1968-1991, and the files represent the universe of 'all deaths in the United States' (www.ICPSR[8]). In preparing the data tapes for this collection and before depositing them in the ICPSR online archives, NCHS protected against potential violations of confidentially by removing direct identifiers and characteristics that could potentially lead to identification of data subjects (www.ICPSR[1]).

Most of the datasets held by ICPSR (including the Mortality Detail Files) can be downloaded direct from the ICPSR website (www.ICPSR[9]) without any restrictions. In most cases, users may access and perform secondary analysis on the micro-data without the need for special legal status or special arrangements (www.NCHS[4]). Only in cases where the nature of the data is considered to be sensitive are restrictions put in place, as in the case of the Community Tracking Study and related surveys, as well as a number of studies acquired in conjunction with the National Archive of Criminal Justice Data (www.NACDJ). In these cases, users are required by ICPSR to complete and sign a Restricted Data Use Agreement, which will allow access to the data for a specific time period (www.ICPSR[3]).

In imposing these non-enforcable obligations, ICPSR requires users to give "assurance that [use] of statistical data will conform to widely-accepted standards of practice and legal restrictions that are intended to protect the confidentiality of research subjects", in terms of the ICPSR Bylaws (www.ICPSR[3]).

It is the principle of "ensuring that federally sponsored and funded data resources are available to all potential users" (www.NCHS[4]), together with the assumption that potential users will conform to "accepted standards of practice", that differentiates the policy regarding access to personal information in the United States from that in England and Wales. The contrast between the mandate to withhold individual-level census-based and vital-events based data from the public domain contrasts dramatically with the US mandate that such information be shared with the public. It is this difference that enabled Stack to conduct studies on occupation-related suicide based on individual-level mortality data (e.g. Stack, 1995; Stack, 1996; Stack, 1998; Stack, 2001) but precludes the possibility of similar studies being conducted in England and Wales, unless they are initiated and conducted within the Office for National Statistics.

In the following section, the different approaches in the United States and England and Wales towards protecting the privacy of data subjects are reviewed.

4. PERSONAL DATA, PRIVACY, AND DATA PROTECTION

The concept of 'privacy' that informs data-protection legislation relates to 'privacy of personal data' (also referred to as 'information privacy'), as opposed to 'privacy of the person', 'privacy of personal behaviour', or 'privacy of personal communications' (www.GO; www.UM[1]). 'Information privacy' refers to the right of the individual to determine when, how, and to what extent, information collected from or about him (or her) (i.e. personal data) for a specific purpose may be conveyed by the information-collector to other parties.

The OECD Guidelines (which are discussed in greater detail later in this section) define *personal data* as "any information relating to an identified or identifiable individual (data subject)" which is held in either the public or private sphere (OECD, 1980). This could include information about a subject's physical or psychological characteristics, family, financial status, cultural or political identity, sexual orientation, racial or ethnic origin, and religious or philosophical convictions. Access to such data could potentially give rise to disclosure of identity of the subject and be construed as invasion of privacy. *Invasion of privacy* involves intrusion of a 'public' actor into the realm of the 'private' without the individual's consent and, in the context of *information privacy* (which is the aspect of privacy that is relevant to the present discussion), is represented by access to personal data (or sharing of personal data) without the subject's consent (Stalder, 2002)

The United States took an early lead in developing policies to protect personal data. In 1973 a task-force was formed at the US Department of Health Education and Welfare (HEW) for the purpose of examining the impact of computerisation on the privacy of medical records. This task-force developed a Code of Fair Information Practices (FIP) which incorporated five principles: openness, disclosure, secondary use, correction and security (www.PRCH). At about the same time, Sweden enacted a law which codified similar fair information principles to those formulated by the HEW (www.PRCH). In the ensuing years, other European countries enacted similar data-protection laws (www.PRCH) and most industrialised nations now have comprehensive privacy and data-protection legislation that extends to both the public and private sectors (www.UM^2). Much of this legislation is based on the Organization for Economic Cooperation and Development's (OECD) *Guidelines on the Protection of Privacy and Transborder Flows of Personal Data*, which was adopted by OECD in 1980 (OECD, 1981). This privacy code was designed to help "harmonize national privacy legislation and, while upholding human rights, [to] at the same time prevent interruptions in international flows of data" (OECD, 1981). By setting out core principles relating to information privacy, the Guidelines represent a consensus between OECD's thirty member-countries (which include the United States and United Kingdom) on basic principles which can be built into national legislation (OECD, 1980).

The Council of Europe's (COE) *Convention for the Protection of Individuals with Regard to Automatic Processing of Personal Data* (COE, 1981), which was adopted in 1981, has strong similarities with the OECD Guidelines, although it applies only to automated data and, by virtue of its membership body, does not apply to the United States. The Council of Europe identified the need for a specific convention to address the specific risks posed by electronic processing, and the primary objective of the Convention is to balance the need to allow the movement of personal data against the need to protect personal privacy (www.UCL).

The COE *Convention* and the OECD *Guidelines* provide much of the conceptual basis for data-protection legislation in Europe and North America. But they operate at a fairly high level of generality and do not mandate specific measures for application of standards or provisions for enforcement amongst member states (www.UM^2), which potentially gives rise to discrepancies in legislative principles between OECD member states, for example the United States and England and Wales (a case which is central to discussion in the present chapter). But the Fair Information Practices that are enshrined in the COE *Convention* and OECD *Guidelines* serve as a minimum standard for data-protection policies and legislation (www.GO). Fair Information Practices (FIPs) are the basic principles for collection, use, disclosure, retention and disposal of personal data and, while there are variations, normally incorporate the principles shown in the table below:

Fair Information Principles
• Ensuring public awareness and transparency of information policies and practices; • Establishing necessity and relevance of the information collected; • Building in finality (i.e. establishing the uses of the information in advance and eventually destroying it); • Identifying the person who has responsibility for protecting personal information within an organization; • Getting informed consent from the individual; • Maintaining accuracy and completeness of records.

Source : www.GO

The OECD *Guidelines* and the COE *Convention* provided the impetus for the introduction of data-protection legislation in England and Wales. The Data Protection Act of 1984 adopted the general principles of the COE *Convention* and the OECD *Guidelines* and built a regulatory framework around it (www.UCL). It was replaced by the Data Protection Act of 1998 which came into force in 2000 and gave effect in UK law to the 1995 EC Data Protection Directive (www.EP). This directive mandated (inter alia) that it should not lead to any diminution of the level of protection already provided in existing national law (www.UCL), and thus reinforces the scope of the Census Act and Population (Statistics) Act in respect of the privacy of census and vital-events data – but only in the case of *living* subjects (www.DCA[1]). The relevance of this proviso to the focus in this thesis on mortality data, relates to discrimination between data held on the Longitudinal Study about living people, and data held about deceased subjects. As the discussion in the present chapter has demonstrated, Longitudinal Study data that relate to living subjects may conceivably be protected by an extra layer of privacy legislation over and above that provided by the Census Act and Population (Statistics) Act, although this cannot be the case with mortality data.

In summary, the differences between the United States and England and Wales in relation to accessibility of national mortality data to the public is *not* connected to differences in privacy legislation, but is accounted for by three other factors, as outlined below:

i) The fundamental curtailment in England and Wales of access to individual-level mortality data, as mandated by the Population (Statistics) Act;
ii) The mandate in the United States that national data (including mortality data) be shared with the public;
iii) The adoption of practices in the United States to anonymise population data so as to allow it to be released to the public while at the same time protecting the privacy of data subjects.

In the following section, the role of the Office of National Statistics in protecting the privacy of Longitudinal Study data subjects, and the implications of this privacy policy for academic research, will be discussed.

5. THE IMPLICATIONS FOR ACADEMIC RESEARCH OF PRIVACY LEGISLATION RELATING TO THE LONGITUDINAL STUDY

The role of the Office for National Statistics (ONS) in managing the privacy of Longitudinal Study subjects is an invidious one. ONS is a government agency which was created in 1996 by a merger of the former Central Statistical Office, the Office for Population Censuses and Surveys (which includes Civil Registration of births, marriages and deaths), and the statistical division of the Department for Employment, thus bringing together a wide range of statistical activities. ONS is charged with collecting, compiling, analysing, disseminating and publishing economic, social, demographic, and labour-market statistics. Its role is to "coordinate activity on official statistics in other government departments, all aimed at giving parliament, government and the wider community the statistical information, analysis and advice they need to improve decision making, stimulate research and inform debate" (www.CS). But, as discussion in earlier sections has demonstrated, the capacity of ONS to fulfill its mission to share with the "wider community" the "statistical information … they need" (www.CS) is fundamentally curtailed, in the case of individual-level data, by the Census Act and the Population (Statistics) Act.

One of the implications of the Census Act and the Population (Statistics) Act is that, since secondary analysis of Longitudinal Data is restricted to officers of ONS who have signed the Official Secrets Act, research initiatives that depend on access to such data, unless initiated and conducted from within, cannot be implemented at all. The only exception to restrictions on direct access by external parties to Longitudinal Study data is in "special cases", such as for research based on coroners' reports, where researchers may (on application) "consult" data, after signing a confidentiality agreement (F. Van Galen, ONS: personal communication, June 2004).

ONS's proposed Integrated Population Statistics System combines census, survey and administrative data, linked at individual-level, to create "a single, comprehensive population statistics database, which is updated over time" (Office for National Statistics, 2003). In envisaging that the database would result in "improved, more consistent statistics for the government community, the Health Service, *academia* and the private sector" (my italics), ONS appears to acknowledge an obligation to the academic community to provide it with data it needs but, at the same time, asserts that "the use of detailed multivariate data … has an associated

risk that information about an identifiable individual will be disclosed" (Office for National Statistics, 2003).

The Census Act and the Population (Statistics) Act place constraints on the Office for National Statistic's potential for sharing Longitudinal Study data with the public and, in the case of individual-level data, the potential for sharing is fundamentally curtailed. This gives rise to a situation whereby research initiatives which depend on access to such data, unless both initiated and conducted from within the Office for National Statistics, cannot be conducted at all. This anomaly demonstrates a situation in England and Wales whereby the trade-off between public good and individual privacy means that legislation may do just as much harm as good. In order to be effective in informing social policy, research is dependent on unobstructed access to population datasets (O'Grady & Nolan, 2004). Legislation which obstructs access to mortality data gives rise to a situation whereby any research initiative that is dependent on access to individual-level mortality detail, is rendered fundamentally unviable.

One of the objectives of the National Centre for Health Statistics is to provide "important surveillance information that helps identify and address critical health problems" (www.NCHS[1]). Legislation in the United States not only facilitates access to population data so as to enable external research, but mandates that such data be made available to the public. By contrast, although the Office for National Statistics aims to "give … parliament, government and *the wider community* the statistical data … they need … to improve decision making, stimulate research and inform debate" (my italics)(www.CS), the Census Act and the Population (Statistics) Act restrict ONS's capacity to fulfill this aim adequately (and, in some cases, even makes research impossible).

The Census Act and Population (Statistics) Act protect the privacy of both the living and the dead. The Data Protection Act, by contrast, protects the privacy of the living. The omission of 'the dead' from its umbrella of protection might be construed as tacit acknowledgement that the dead do not need protection. This observation prompts the question: To what extent *do* the dead need protection? This is a question that can neither be addressed nor answered adequately within the scope of this thesis, but one which should be central to any future discussion about the implications of the Census Act and the Population (Statistics) Act for research relating to mortality.

6. CONCLUSION

Restrictions on access to relevant mortality data proved to be the factor that decided the feasibility of the research proposal outlined in this thesis. Even though the proposal, as it stood, ultimately proved to be unviable, the literature review that is central to this thesis highlighted a series of important issues relevant to study of

the relationship between occupation and suicide that could serve as the basis for future investigation. But, as discussion has shown, such future investigations must necessarily be based on the assumption that access to individual-level mortality data is not possible in England and Wales – unless a source of such data exists which escaped my attention during the search for a source of primary mortality data as outlined in Chapter 4.

One of the implications of both the Census Act and the Population (Statistics) Act is that research initiatives that depend on access to individual-level mortality data (and, indeed individual-level data of any other category) held on the Longitudinal Study can only be initiated from within *and* conducted within the Office for National Statistics, by authorised members of ONS staff. This constraint on mortality-related social-scientific enquiry stands by stark contrast to the situation in the United States, where population data are readily accessible from both the National Centre for Health Statistics and the Inter-University Consortium for Social and Political Research. This contrast brings a series of questions to the fore:

i) To what extent is it reasonable to protect the privacy of the dead, at the expense of valid social-scientific enquiry?
ii) Are some of the provisions of the Census Act and Population (Statistics) Act (and, in particular, those provisions that restrict the sharing of population data with the academic research community) outdated? Could these Statutes justifiably be updated to facilitate social-scientific research designed to inform social policy in a *contemporary* society?
iii) To what extent and by what legislation could England and Wales emulate the United States' practice of anonymising national data so that it can be released to the public while at the same time the privacy of data subjects is protected?

For the meantime, however, research into the suicide-occupation relationship must be based on the assumption that the Longitudinal Study is a futile source of individual-level mortality data as well as an inadequate source of aggregate mortality data (as discussed in Chapter 5, section 5).

Examining the overlap between the findings of two studies (Stack, 2001; Kelly & Bunting, 1998) that are central to discussion in this thesis yielded insights into how the different theoretical premises that underlie analysis of the suicide-occupation relationship – and the different methodologies these premises inform – gives rise to some results that are consistent between the studies and some that are discrepant. The results of these studies are similar in indicating medical practitioners, dental practitioners, and nurses to be at significantly high risk of suicide. This outcome is not surprising, despite the discrepant theoretical premises and methodologies that underlie the studies, since similar findings have been

demonstrated by many other studies (Kelly & Bunting, 1998)(also see Chapter 5, section 2).

Explanation for the high suicide risk associated with these occupations tends to be biased towards ease of access to drugs and expertise in their use (e.g. Hawton et al, 2000). Such an explanation is consistent with the concept of suicide risk being associated with 'opportunity factors' (as discussed in Chapter 2, section 3.3), a contributory factor to suicide risk which is incorporated into Stack's *multicausal heuristic model of occupation-related suicide* (as detailed in Chapter 3, section 2).

While it is true that doctors, dentists and nurses have easier access to drugs than do many other occupations, my review of the theoretical literature did not reveal any evidence that the relationship between ease of access to drugs and the significantly high rates of suicide for these occupations is necessarily a causal one. My own hunch (as discussed in Chapter 4, section 3) is that the level of control (or lack of) that a practitioner has over the conditions under which the occupation is practiced might be an equally significant contributory factor to the associated suicide risk. This explanation would be consistent, in terms of Stack's model, with the concept of suicide risk being associated with 'stress internal to the nature of the occupation itself' (see Chapter 2, section 3.1).

'Occupational control' may be one aspect of 'stress internal to the nature of the occupation itself' and refers (in terms of my discussion in Chapter 4, section 3) to the degree to which the practice of an occupation is controlled by institutional and structural factors, such as central government directives and legislation concerning working practices. Control mechanisms (e.g. government inspections) mandated by these institutions tend to introduce an extra layer of stress over and above that already associated with practice of that occupation (e.g. exposure to death).

Developing a scale to measure 'occupational control' would allow this hunch to be tested, by enabling the level of 'occupational control' associated with a range of occupations to be computed and compared. By plotting the 'occupational control' indicators for a range of occupations against relative suicide-risk indicators, future research could investigate the influence of 'occupational control' as a predictor for occupational suicide rates. This exercise could, in the case of doctors, dentists and nurses, provide a counterbalance to the tendency to bias explanations for high suicide rates towards 'opportunity factors'.

The conditions under which police-officers work bear significant commonalities to those under which medical personnel work, particularly in relation to the potential for exposure to traumatic situations such as motor accidents and deaths and (in the case of police-officers) discharge of firearms and assault. In terms of the 'stress internal to the nature of the occupation itself' dimension of Stack's *multicausal heuristic model of occupation-related suicide*, the conditions under which police-officers work constitute a significant predictor factor for suicide. And yet, by contrast with the findings for medical personnel,

research on police suicide is marked by inconsistent findings, with some studies reporting a 'high' rate of suicide, some an 'average' rate and some a 'low' rate (e.g. Stack & Kelley, 1994) (as discussed in Chapter 2, section 2).

My own work (in the Occupational Health and Welfare Unit of Greater Manchester Police) suggests that certain features of working in a police-force characteristically differentiate policing from medical practice. One of the most significant of these is, in my view, the culture of 'camaraderie' that typifies policing, a culture that may be less evident in medical practice, nursing and dentistry. My hunch is that this 'camaraderie' mitigates to a considerable degree against the high levels of 'stress internal to the nature of the occupation itself' (Stack, 2001), and differences in measures of 'camaraderie' demonstrated by different police-forces may be a factor that contributes to discrepant suicide-risk findings. This 'hunch' is one that warrants further investigation.

Measures of suicide risk (both within the general population and for individual occupations or occupational groups) are precarious over time, because of such influences as longitudinal trends in suicide in the general population, changes in fashions in suicide, and demographic shifts in the general population and within occupations themselves (as discussed in Chapters 2 and 3). This means that there cannot be a definitive and stable value for the suicide risk associated with any particular occupation. This precariousness tends to confound the meaningfulness of direct comparisons of suicide statistics that are based on data from different time periods. But, even as a 'snapshot' in time, measures of suicide risk must be a tentative estimate only. The reliability and validity of estimates are difficult to establish for a variety of reasons, many of which have been highlighted in this thesis. The first of these relates to the use of different measures for evaluating suicide risk. In the example demonstrated in this thesis, the Provisional Mortality Ratio is used as a measure of suicide risk for two local studies (Kelly et al, 1995; Kelly & Bunting, 1998), while multivariate logistic regression odds ratios are used for an American study (Stack, 2001). Since the theoretical assumptions that underlie use of these measures differ, the indicators are not directly comparable with each other.

Comparison between studies is also complicated by the use of different protocols for coding occupations. The Classification of Occupations provides the basis for occupational classification in England and Wales, whereas the American study central to this thesis (Stack, 2001) is based on standard federal occupational codes. Similarly, conventions for defining 'suicide' are complex (as discussed in Chapter 2, Section 5.1) and differ between studies, between countries, and between historical eras. This anomaly introduces an extra complexity into any analysis of suicide risk that has a comparative dimension.

As well as highlighting the complexities involved in analysis of suicide risk, this thesis has demonstrated the significance of controlling for the socio-demographic variables of occupations (in particular, age, gender, marital status and

race), with specific reference to a comparative analysis of a local study (Kelly & Bunting, 1998) and an American study (Stack, 2001) (as discussed in 3). The theoretical assumptions and methodology that are central to Stack's study provided the basis for the research proposal outlined in this thesis, which sought to answer the following questions:

i) Are there occupational groups in England and Wales that have a significantly high or significantly low risk of suicide in relation to comparable groups in the general population?
ii) How do the demographic variables of age, sex, marital status and ethnicity mediate the risk of occupation-related suicide (remember to adjust actual application in this section)?
iii) Does the process of introducing controls for the demographic covariates of occupation support or refute the findings of Kelly & Bunting's (1998) study?

This proposal ultimately proved to be unviable, but this thesis has identified a series of legislative differences between England and Wales and the United States that account for the differences in accessibility of national mortality data to the public. These factors are itemized as follows:

i) The fundamental curtailment (in England and Wales) of access to population data (including individual-level mortality data) that is mandated by the Census Act and the Population (Statistics) Act;
ii) The mandate in the United States that national data, including mortality data, be shared with the public;
iii) The adoption of practices in the United States to anonymise national data so as to allow it to be released to the public while at the same time protecting the privacy of data subjects.

This thesis has identified a series of important considerations that should be taken into account in any analysis of occupation-related suicide and has highlighted a number of possibilities for future research initiatives. But it has also identified a question that is fundamental to further discussion of the implications of the Census Act and the Population (Statistics) Act for research and public policy relating to mortality:

To what extent do the dead need protection?

REFERENCES

Note : Websites and web-pages are listed separately.

Adelstein, A. & Mardon, C. (1975) Suicides 1961-74. *Population Trends*, 3:13-18.

Bennett, C. & Grant, R. (eds). (1999) *Visions of Privacy*. Toronto: University of Toronto Press.

Bar, O, Pahlke, C, Dahm, P, Weiss, U. & Heuft, G. (2004) Secondary prevention for police officers involved in job-related psychological stressful or traumatic situations. *Z Psychosom. Med Psychother*, 50(2):190-202. German.

Boxer, P, Burnett, C. & Swanson, N. (1995) Suicide and occupation: a review of the literature. *Journal of Occupational and Environmental Medicine,* 37:442-52.

Brin, D. (1998) *The transparent society: will technology force us to choose between privacy and freedom?* Reading, MA: Perseus Books

British Union of Social Work Employees (BUSWE). 2004. Stress in the workplace. *The Quarterly Bulletin for BUSWE members*. March 2004, p.1.

Bulusu, L. & Alderson, M. (1984) Suicides 1950 – 82. *Population Trends*, 35:11-17.

Bunting, J. & Kelly, S. (1998) Geographic variations in suicide mortality, 1982-96. *Population Trends,* 93:7-18.

Charlton, J, Kelly, S, Dunnell, K, Evans, B, Jenkins, R. & Wallace, R. (1992) Trends in suicide deaths in England and Wales. *Population Trends,* 69:10-16.

Charlton, J. (1995) Trends and patterns in suicide in England and Wales. *International Journal of Epidemiology*, 24: S45-S52.

Charlton, J. & Murphy, M. (eds). (1997) *The health of adult Britain: 1841-1994*. London: Stationery Office.

Council of Europe. 1981. *Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data*. European Treaty Series, No. 108.

De la Monte, S.M. & Hutchins, G.M. (1984) Is suicide a special occupational hazard for physicians? *Journal of the American Medical Association*, 251(15):1952-3.

Drever, F. (ed.) (1995). *Occupational health - decennial supplement*. Office for National Statistics. London: The Stationery Office. Series DS, No. 10.

Drever, F. & Whitehead, M. (eds). (1997) *Health Inequalities*. Office for National Statistics. London: The Stationery Office. Series DS, No. 15.

Drever, F. & Bunting, J. (1997) Patterns and trends in male mortality. IN: Drever, F. and Whitehead, M. (eds) *Health inequalities*. London: The Stationery Office.

Durkheim, E. (1952) *Suicide: a study in sociology*. Translated by John A. Spaulding and George Simpson. London: Routledge and Kegan Paul.

Gibbs, J.P. & Martin, W.T. (1958) A theory of status integration and its relationship to suicide. *American Sociological Review*, 23(2):140-147.

Griffiths, C. & Brock, A. (2003) Twentieth century mortality trends in England and Wales. *Health Statistics Quarterly*, 18:5-17.

Hattersley, L. & Creeser, R. (1995) *Longitudinal study 1971-1991: history, organisation and quality of data*. London: HMSO.

Hawton, K. (1987) Assessment of suicide risk. *British Journal of Psychiatry*, 150:145-53.

Hawton, K. & Vislisel, L. (1999) Suicide in nurses. *Suicide and Life Threatening Behaviour*, 29 (1):86-95.

Hawton, K. Fagg, J. Simkin, S. Harriss, L. & Malmberg, A. (1998) Methods used for suicide by farmers in England and Wales. *British Journal of Psychiatry*, 173:320-4.

Hawton, K, Ware, C, Mistry, H, Hewitt, J, Kingsbury, S, Roberts, D. & Weitzel, H. (1995) Why patients choose paracetamol for self-poisoning and their knowledge of its dangers. *British Medical Journal,* 310: 164.

Hawton, K, Clements, A, Simkin, S & Malmberg. A. (2000) Doctors who kill themselves: a study of the methods used for suicide. *Quarterly Journal of Medicine*, 93:352-7.

Hood-Williams, J. (1996) Studying suicide. *Health & Place*, 2(3):167-177.

Jick, S, Dean, A. & Jick, H. (1995) Antidepressants and suicide. *British Medical Journal,* 310: 215-18.

Kelly, S, Charlton, J, & Jenkins, J. (1995) Suicide deaths in England and Wales, 1982–92: the contribution of occupation and geography. *Population Trends,* 80:16–25.

Kelly, S. & Bunting, J. (1998) Trends in suicide in England and Wales, 1982–96. *Population Trends,* 92:29 - 41.

Kushner, H. (1993) Suicide, gender, and the heart of modernity in nineteenth century medical and social thought. *Journal of Social History*, 26:461-90.

Labovitz, S. & Hagedorn, R. (1971) An analysis of suicide rates among occupational categories. *Sociological Inquiry,* 41:67-71.

Lampert, D.I, Bourque, L.B. & Kraus, J.F. (1984) Occupational status and suicide. *Suicide and Life Threatening Behaviour,* 14(4)254-69.

Lyon, David. (1994) *The Electronic Eye: The Rise of the Surveillance Society*. Minneapolis: University of Minnesota Press.

Malmberg, A. Simkin, S. & Hawton, K. (1999) Suicide in farmers. *British Journal of Psychiatry,* 175:585-6.

Milham, S. (1983) *Occupational mortality in Washington State*. Olympia: Washington State Department of Social and Health Services.

O'Grady, K.F. & Nolan, T.M. (2004) Privacy: bad for your health? *Medical Journal of Australia,* 180(6):307; author reply.

Office for National Statistics (ONS). (2003) *Proposals for an integrated population statistics system*. Office for National Statistics. Discussion paper, October 2003. See: www.statistics.gov.uk/downloads/theme_population/ipss.pdf

Organization for Economic Cooperation and Development (OECD). 1981. *Guidelines to the Protection of Privacy and Transborder Flows of Personal Data.* Paris: OECD.

Peloso, P.F. (1992) Approach to the study of suicide in the essays of Enrico Morselli (1879) and in the successive hypotheses of Durkheim and Freud. *Pathologica*, 84(1089):107-20. Italian.

Powell, E.H. (1958) Occupation, status, and suicide: toward a redefinition of anomie. *American Sociological Review*, 23(2)131-139.

Price, D.H. (1997) Anthropological Research and the Freedom of Information Act. *Cultural Anthropology Methods,* 9(1):12-15.

Rees, P, Martin, D. & Williamson, P. (2002) Census data resources in the United Kingdom. IN: Rees, P., Martin, D. and Williamson, P. (eds) *The census data system.* Chichester: Wiley, p.1-24.

Rooney, C. & Devis, T. (1996) Mortality trends by cause of death in England and Wales 1980-94: the impact of introducing automated cause coding and related changes in 1993. *Population Trends,* 86:29-35.

Rooney, C. & Smith, S.K. (2000). Implementation of ICD-10 for mortality data in England and Wales from January 2001. *Health Statistics Quarterly,* 8:41-50

Stack, S. & Kelley, T. (1994) Police suicide: an analysis. *American Journal of Police,* 13(4):73–90.

Stack, S. (1995) Suicide risk among laborers: a multivariate analysis. *Sociological Focus,* 28(2):197-99

Stack, S. (1996) Suicide risk among dentists: a multivariate analysis. *Deviant Behaviour,* 17:107-17.

Stack, S. (1997) Suicide among artists. *Journal of Social Psychology*, 137(1):129-30.

Stack, S. (1998) *Suicide risk among physicians: a multivariate analysis*. Paper presented at the annual meetings of the American Association of Suicidology, April 2000, Bethesda, Maryland.

Stack, S. (2001) Occupation and suicide. *Social Science Quarterly*, 82 (2):384-396.

Stalder, F. (2002) The failure of privacy enhancing technologies (PETS) and the voiding of privacy. *Sociological Research Online,* 7(2). www.socresonline.org.uk/7/2/stalder.html

Taylor, S. (1982) *Durkheim and the study of suicide*. London: McMillan.

Travis, R. (1990) Halbwachs and Durkheim: a test of two theories of suicide. *British Journal of Sociology,* 41(2): 225-43.

Violanti, J.M, Vena, J.E. & Marshall, J.R. (1996) Suicides, homicides, and accidental death: a comparative risk assessment of police officers and municipal workers. *American Journal of Industrial Medicine*, 30(1):99-104.

Wenz, F.V. (1977) Effects of seasons and sociological variables on suicidal behaviour. *Public Health Rep*, 92(3):233-9

World Health Organization (WHO). 1977. *International statistical classification of diseases, injuries and causes of death (ICD-9)*, 9th revision. Vols. 1-2. Geneva: WHO.

World Health Organization (WHO). 1992-1994. *International statistical classification of diseases and related health problems (ICD-10)*. 10th revision. Vols. 1-3. Geneva: WHO.

The listed websites and web-pages were all accessed at some (or several) point(s) in time between November 2003 and June 2004.

www.BBC — BBC News, World Edition. *Profile: Dr David Kelly*. 27 January, 2004
http://news.bbc.co.uk/2/hi/uk_news/politics/3076869.stm

www.CDC[1] — Centre for Diseases Control and Prevention (CDC).
www.cdc.gov

www.CDC[2] — Centre for Diseases Control and Prevention (CDC).
Self inflicted injury / suicide.
www.cdc.gov/nchs/fastats/suicide.htm

www.CDC[3] — Centre for Diseases Control and Prevention (CDC).
Downloadable public-use data files and documentation from the Inter-university Consortium for Political and Social Research (ICPSR).
www.cdc.gov/nchs/datawh/icpsr.htm)

www.CeLSIUS[1] — Centre for Longitudinal Study Information and User Support (CeLSIUS)
www.celsius.lshtm.ac.uk

www.CeLSIUS[2] — Centre for Longitudinal Study Information and User Support (CeLSIUS)
How can I use the ONS Longitudinal Study?
www.celsius.lshtm.ac.uk/how.html

www.CeLSIUS[3] — Centre for Longitudinal Study Information and User Support (CeLSIUS)
Step by step guide to using the ONS Longitudinal Study.
www.celsius.lshtm.ac.uk/stepbystep.html

www.CeLSIUS[4] — Centre for Longitudinal Study Information and User Support (CeLSIUS)
LS outputs: changes in coding schemes.
www.celsius.lshtm.ac.uk/modules/outputs/op030200.html

www.CeLSIUS[5] Centre for Longitudinal Study Information and User Support (CeLSIUS)
LS outputs: introduction.
www.celsius.lshtm.ac.uk/modules/outputs/op010000.html

www.CeLSIUS[6] Centre for Longitudinal Study Information and User Support (CeLSIUS)
LS outputs: what can be published.
www.celsius.lshtm.ac.uk/modules/outputs/op040100.html

www.CeLSIUS[7] Centre for Longitudinal Study Information and User Support (CeLSIUS)
LS outputs: aggregation.
www.celsius.lshtm.ac.uk/modules/outputs/op040200.html

www.CS Civil Service Recruitment Gateway.
Office for National Statistics.
www.careers.civil-service.gov.uk/Index.asp?txtNavID=
124&txtOverRideDocID=165&635132

www.DCA[1] Department for Constitutional Affairs (DCA).
Data protection – frequently asked questions.
www.dca.gov.uk/ccpd/faqdp.htm)

www.DCA[2] Department for Constitutional Affairs (DCA).
Data sharing.
www.dca.gov.uk/foi/sharing/index.htm).

www.DU Dundee University (DU). Forensic medicine unit.
Medico-legal systems of death investigation.
www.dundee.ac.uk/forensicmedicine/llb/deathinvestig.htm

www.DPHA Dyfed Powys Health Authority.
Director of public health, annual report 2001.
www.dyfpws-ha.wales.nhs.uk/dirphanrep2001/contents.htm

www.GO Government of Ontario.
Privacy impact assessment guidelines.
www.gov.on.ca/mbs/english/fip/pia/pia2.html

www.HSE Health and Safety Executive (HSE).
www.hse.gov.uk

www.HHS[1] United States Department of Health and Human Services (HSS).
www.hhs.gov

www.HHS[2] U.S. Department of Health and Human Services (HHS).
Current legislative authorities of the National Centre for Health Statistics.
www.cdc.gov/nchs/data/misc/legis99.pdf

www.HMD Human Mortality Database (HMD).
www.mortality.org

www.ICPSR[1]	Inter-University Consortium for Political and Social Research (ICPSR). www.icpsr.umich.edu
www.ICPSR[2]	Inter-University Consortium for Political and Social Research (ICPSR). *Description – study no. 7632, Mortality Detail Files, 1968-1991.* http://webapp.icpsr.umich.edu/cocoon/ICPSR-STUDY/07632.xml
www.ICPSR[3]	Inter-university Consortium for Political and Social Research (ICPSR). *Restricted data.* www.icpsr.umich.edu/access/restricted/index.html
www.ICPSR[4]	Inter-university Consortium for Political and Social Research (ICPSR). *Data use tutorial.* www.icpsr.umich.edu/help/newuser.html#00
www.ICPSR[5]	Inter-university Consortium for Political and Social Research (ICPSR). *General information about restricted data from ICPSR.* www.icpsr.umich.edu/access/restricted/general.html
www.ICPSR[6]	Inter-university Consortium for Political and Social Research (ICPSR). *Description--Study No. 7632: mortality detail files 1968-1991.* webapp.icpsr.umich.edu/cocoon/ICPSR-STUDY/07632.xml
www.ICPSR[7]	Inter-university Consortium for Political and Social Research (ICPSR). *Downloads – study no. 7632.* www.icpsr.umich.edu/cgi/archive.prl?study=7632)
www.ICPSR[8]	Inter-university Consortium for Political and Social Research (ICPSR). *Description--Study No. 7632: mortality detail files 1968-1991.* http://webapp.icpsr.umich.edu/cocoon/ICPSR-STUDY/07632.xml
www.ICPSR[9]	Inter-university Consortium for Political and Social Research (ICPSR). *Data access and analysis.* www.icpsr.umich.edu/access/index.html
www.NACDJ	National Archive of Criminal Justice Data (NACDJ) *Restricted access data archive.* www.icpsr.umich.edu/NACJD/Private/index.html#listing.
www. NCHS[1]	National Centre for Health Statistics (NCHS). www.cdc.gov/nchs/
www. NCHS[2]	National Centre for Health Statistics (NCHS). *International Statistical Classification of Diseases, Ninth Revision (ICD-9.* www.cdc.gov/nchs/about/major/dvs/icd9des.htm#rev
www. NCHS[3]	National Centre for Health Statistics (NCHS). *International Statistical Classification of Diseases, Tenth Revision (ICD-10).* www.cdc.gov/nchs/about/major/dvs/icd10des.htm

www.NCHS[4] National Centre for Health Statistics (NCHS). *Policy on micro-data dissemination.* www.cdc.gov/nchs/data/NCHS%20Micro-Data%20Release%20Policy%204-02A.pdf

www.NDAD[1] National Digital Archive of Datasets (NDAD). http://ndad.ulcc.ac.uk

www.NDAD[2] National Digital Archive of Datasets (NDAD). *Statistical departments: historic mortality data files, series catalogue CRDA/20* http://ndad.ulcc.ac.uk/CRDA/20/detail.html

www.NDAD[3] National Digital Archive of Datasets (NDAD). *Historic Mortality Data Files, 1901-1992 dataset. NDAD Reference: CRDA/20/DS/1* http://ndad.ulcc.ac.uk/CRDA/20/DS/a/summary.html

www.NDAD[4] National Digital Archive of Datasets (NDAD). *Historic Mortality Data Files, 1901-1995 dataset. NDAD Reference: CRDA/20/DS/2* http://ndad.ulcc.ac.uk/CRDA/20/detail.html

www.NDAD[5] National Digital Archive of Datasets (NDAD). *Statistical departments: historic mortality data files, series catalogue CRDA/20: identity statement* http://ndad.ulcc.ac.uk/CRDA/20/detail.html#identity_statement

www.NHS National Health Service (NHS) Electronic Library for Health. *Suicide.* http://cebmh.warne.ox.ac.uk/cebmh/elmh/nelmh/suicide/guide/coroner.html

www.ONS[1] Office for National Statistics (ONS). *Twentieth century mortality (England and Wales, 1901-2000).* CD-ROM. www.statistics.gov.uk/StatBase/Product.asp?vlnk=1253&More=Y

www.ONS[2] Office for National Statistics (ONS). *The ONS Longitudinal Study.* www.statistics.gov.uk/StatBase/Source.asp?vlnk=904&More=Y

www.ONS[3] Office for National Statistics (ONS). *Discussion paper: proposals for an integrated population statistics system*, 2003. www.statistics.gov.uk/downloads/them_population/ipss.pdf

www.ONS[4] Office for National Statistics (ONS). *Comparability Ratios for major causes of death, by sex, England and Wales.* www.statistics.gov.uk/about/classifications/icd10/comparability_ratios.asp

www.ONS[5] Office for National Statistics (ONS). *ICD-10 for mortality.* www.statistics.gov.uk/about/classifications/icd10/default.asp.

www.ONS[6] Office for National Statistics (ONS).
National Statistics Online, homepage.
www.statistics.gov.uk/default.asp

www.ONS[7] Office for National Statistics (ONS).
A Demographics Statistics Service for the 21st Century: Introductory covering letter from Len Cook, National Statistician and Registrar General for England & Wales
www.statistics.gov.uk/about/Methodology_by_theme/Dem_Stat_Ser_21ST_Cen.asp

www.ONS[8] Office for National Statistics (ONS).
ONS Initiatives: Proposals for population statistics.
www.statistics.gov.uk/CCI/nugget.asp?ID=531&Pos=6&ColRank=2&Rank=224

www.PRCH Privacy Rights Clearing House (PRCH).
Fair information principles.
www.privacyrights.org/ar/fairinfor.htm

www.TU Trent University (TU).
Sociology workshop.
www.trentu.ca/academic/sociology/soci210/durkwork.htm

www.UCL University College of London, Constitution Unit.
Data Protection: a brief history.
www.ucl.ac.uk/constitution-unit/foidp/dp_history.php

www.UKDA[1] UK Data Archive (UKDA).
www.data-archive.ac.uk

www.UKDA[2] UK Data Archive (UKDA).
SN2902 : Historic mortality and population data, 1901-1992
www.data-archive.ac.uk/findingData/snDescription.asp?sn=2902

www.UKDA[3] UK Data Archive (UKDA).
SN3625: Local mortality datapack – population and deaths by cause, 1979-1992
www.data-archive.ac.uk/findingData/snDescription.asp?sn=3625

www.UM[1] University of Miami. Privacy / data-protection project.
http://privacy.med.miami.edu/glossary/xd_privacy_basicdef.htm

www.UM[2] University of Miami. Privacy / data-protection project.
http://privacy.med.miami.edu/glossary/xd_international_dp_conv.htm

www.PHSA United States Food and Drug Administration.
Public Health Service Act.
www.fda.gov/opacom/laws/phsvcact/phsvcact.htm

www.WHO World Health Organization (WHO).
Mortality Database.
www3.who.int/whosis/menu.cfm?path=whosis,mort&language=english

APPENDIX

APPLICATION TO USE DATA FROM THE ONS LONGITUDINAL STUDY FOR CONSIDERATION BY THE LS RESEARCH BOARD

FORM A: User and Study Details and Undertaking

Study Title: Occupation-related suicide:
The mediating impact of demographic co-variables

SECTION A1 PROJECT STAFF DETAILS

Head of proposed study:

Name:	Laurel Milwid
Position:	Postgraduate student – M.Res. (Sociology)
Organisation:	Institute for Social, Cultural and Policy Research, University of Salford
Address:	Crescent Road, Salford, Greater Manchester M5 4WT
Telephone:	0161 295 5000
E-mail:	l.m.milwid@pgr.salford.ac.uk

Collaborators / research staff / supervisors:

Name:	Dr G.W. Smith	Dr N. Hazel
Telephone:	0161 295 5000	0161 295 5000
E-mail:	g.w.h.smith@salford.ac.uk	n.hazel@salford.ac.uk

Expected duration of project

From:	March 2004	**To:**	August 2004

Source of funding:

Tuition fees

SECTION A2 BACKGROUND AND AIMS OF STUDY

1. Give an account of the background to the study

BACKGROUND

Stack (2001) argues that "the rationale for differences in suicide by occupation remains unexplored in any rigorous sense", and that research on occupations and suicide is often marked by inconsistent findings. For example, eleven of eighteen studies of police suicide report a "high" rate, three report an "average" rate, and four a "low" rate (Stack & Kelly, 1994).

Stack suggests that methodological shortfalls in previous studies contribute to such discrepant findings. He points out that research on occupation-related suicide has tended to neglect multivariate models and, as a result, it is not clear whether people in allegedly "high-risk" occupations are at risk of suicide because of occupational stress associated with that occupation or because of the demographic composition of the people in that occupation.

To address these shortcomings, Stack employs a logistic regression model of occupation and suicide with controls introduced for basic demographic correlates of occupation: gender, race, age, and marital status. This model allows the risk of suicide in an occupation to be calculated independent of the demographic covariates of that occupation.

Stack investigates the relationship between occupation and suicide for 32 occupational groups. Although, according to Stack, bivariate logistic regression models find a total of 15 occupations in the United States with either significantly higher or significantly lower risk than the rest of the working-age population, Stack's multivariate model, in removing the demographic covariates of occupation, found only 8 occupations to be either at higher (including dentists, artists, auto-mechanics, and carpenters) or lower (including clerks, elementary school teachers, and cooks) risk of suicide than the expected risk in the rest of the working-age population.

Stack's findings underscore the importance of controlling for the demographic covariates of occupations, and point to the apparent paucity in the United Kingdom of research that recognizes the significance of controlling for the demographic covariates of occupation when analysing suicide risk. A study by Charlton (1995) recognised the importance of including demographic controls in the analysis of suicide risk and occupation, but later studies, including that of Kelly & Bunting (1996), have neglected to incorporate demographic controls.

RESEARCH PROBLEM

The questions that I would be seeking to investigate are:

i. Are there occupational groups in the UK that have significantly high or significantly low risk of suicide?
ii. How do the demographic variables of age, gender, marital status, ethnicity, immigration status, socioeconomic status, geography, and social class mediate the risk of occupation-related suicide?
iii. Does the process of introducing controls for the demographic covariates of occupation support or refute the findings of Kelly & Bunting's study (which identified approximately 32 high-risk occupations for men and 18 for women)?

REFERENCES

Charlton, J. 1995. Trends and patterns in suicide in England & Wales. *International Journal of Epidemiology,* 24: S45-S52.

Kelly, S, & Bunting, J. 1996. Trends in suicide in England & Wales, 1982-96. *Population Trends*, 92: 29-41.

Stack, S. 2001. Occupation and suicide. *Social Science Quarterly*, 82(2): 384-396.

Stack, S, & Kelly, T. 1994. Police suicide: an analysis. *American Journal of Police*, 13(4): 73-90.

2. How will the LS contribute to the aims of the study?

PURPOSE OF THE STUDY

The purpose of the study is to test Stack's findings, using local data, and thus to address an (apparently) neglected area of research in occupation-related suicide. My search for the individual-level data that would enable me to perform logistic regression analysis has shown that getting direct access to such data is not possible in the United Kingdom. It is this search that has led to my correspondence with staff at CeLSIUS and ONS and my application for assistance with my project.

SAMPLES

The project will be based on two samples: an experimental group and a control group.

The *experimental group* will comprise all suicides since the inception of the Longitudinal Study and (based on calculations made by your Alec Ross) will number in the region of 1500 (approx. 60 per year x 30).

The *control group* will comprise all other deaths in the same period. (I am not sure yet whether this sample should perhaps include only *natural* deaths.)

CeLSIUS's contribution would be to derive these samples.

OCCUPATIONAL GROUPS

Stack's study is based on 23 occupational groups. ONS staff have explained the complications concerning violations of confidentiality that could potentially arise by splitting the small number of relevant suicide deaths by occupational group. I am working on the coding scheme that would permit comparison with Stack's US study and, in doing this, will bear in mind the consideration relating to numbers and confidentiality.

CeLSIUS's contribution would be to carry out the re-coding on the basis of my specifications.

VARIABLES

The variables that I anticipate introducing into the analysis are age, gender, ethnicity, geography, marital status, social class, and immigration, as these are variables shown in earlier studies in the UK to be related to suicide risk (Kelly & Bunting, 1996).

CeLSIUS's contribution would be to perform the analyses, on the basis of my specifications.

LOGISTIC REGRESSION ANALYSIS

CeLSIUS's contribution would be to perform the analyses, on the basis of my specifications. This would require ongoing collaboration in order to secure meaningful results.

SECTION A3 PUBLICATION AND DISSEMINATION OF LS DATA

Please indicate how you will publicise and disseminate the findings of your proposed study.

Master's thesis.

SECTION A4 UNDERTAKING

I have read and agree to the conditions relating to the use and publication of the Longitudinal Study and recognize that any breach of confidentiality may result in legal action.
I will seek the agreement of ONS for any extension of the project in terms of substance or duration.

Disclosure of information on individuals

Papers, abstracts, etc. will not identify individuals or enable individuals to be identified. I will not use tables, etc. supplied to me to attempt to obtain or derive information relating specifically to an individual or household, nor claim to have obtained or derived such information. I will not allow tables supplied to me to be used by people outside the approved project without the written permission of ONS.

Retention of tables

If requested to do so by ONS, I will destroy or return to ONS/CeLSIUS all copies of LS tabulations supplied to me.

Publicising details of the study

I understand that, if my application is accepted, information on the project and project staff will be kept on computer files held by CeLSIUS.

Use of newly-derived tables

I understand that variables derived in the course of preparing the data may be added to the LS data files and made available to other users.
I agree to my name and contact details being added to the central list of researchers using services funded under the ESRC/JISC Census of Population Programme.

Conference/seminar papers and papers submitted to journals.

If it is intended to present any unpublished LS data at a conference/seminar, an abstract is required for clearance by ONS. Any draft paper using LS data must be cleared with ONS before submission to a journal. Twenty eight days' notice is required for clearance in both cases. All outputs should acknowledge ONS/CeLSIUS and include a conventional disclaimer. Researchers who are based in the academic community and who have received assistance from CeLSIUS, LSHTM, should send abstracts/papers to Jo Tomlinson at CeLSIUS.

Signed: *L. M. Milwid* 17.04.2004

Academic users should e-mail this form to celsius@lshtm.ac.uk

Also send a signed copy to: **CeLSIUS**
Centre for Population Studies
London School of Hygiene & Tropical Medicine
49-51 Bedford Square
London WC1B 3DP