Most UK datasets of routinely collected health statistics fail to collect information on ethnicity and religion

Citation for published version:

Digital Object Identifier (DOI):
10.1258/jrsm.2008.080007

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Journal of the Royal Society of Medicine

Publisher Rights Statement:
Copyright © 2008, The Royal Society of Medicine

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Most UK datasets of routinely collected health statistics fail to collect information on ethnicity and religion

Kishwar Sultana1 • Aziz Sheikh2
1 General Practitioner, Glasgow
2 Professor of Primary Care Research and Development, Division of Community Health Sciences: GP Section, University of Edinburgh, Edinburgh EH8 9DX, UK
Correspondence to: Aziz Sheikh. E-mail: Aziz.Sheikh@ed.ac.uk

Introduction
The reduction of ethnic inequalities is a long-declared UK government priority,1 but despite the moral and increasingly legal imperative to provide equitable healthcare to all sections of the population, there is very limited evidence of progress in achieving this objective. More recently, New Labour has also committed itself to tackling the very considerable religious inequalities in health and social outcomes that have become evident from analysis of data from the 2001 Census.2 Similar ethnic and religious inequalities in health outcomes almost certainly exist in many other pluralist societies.

Given the difficulties in reducing health inequalities for certain disorders, the very considerable gaps remaining in our knowledge in relation to minority communities for many other conditions, and the known under-representation of minority groups in research (both in the UK and US),3,4 it is important that every effort is made to make use of existing data sources to describe and understand the nature of ethnic- and faith-based variations in health outcomes, and assess progress in tackling these inequalities.

The UK enjoys some of the foremost datasets of routinely collected health statistics, and greater exploitation of these is potentially of considerable importance to shaping policy, prioritizing research and identifying foci for service delivery improvements. In order to investigate the fitness for purpose of these datasets, we sought to interrogate them for evidence of inclusion of ethnicity and faith variables and, where recorded, to see whether there was a consistent approach to recording that would allow comparisons between datasets.

Methods
We interrogated the Directory of Clinical Databases (DoCDat), which is a comprehensive, freely available UK compilation of 162 local and national health datasets. These datasets contain records of demographic and clinical data from individuals presenting to different health-care providers or participating in academic studies. Such clinical and/or research encounters provide an important opportunity to obtain data that could be used to assess disease profiles, health services’ use and clinical outcomes across and between ethnic and faith groupings.

Launched in 1999, DoCDat was developed and has since been maintained by the London School of Hygiene and Tropical Medicine.5 A structured questionnaire is used by those compiling and maintaining DoCDat to determine general details about individual datasets relating to, for example, when the dataset was established and what fields it contains. The quality of datasets is also assessed by scrutinizing the validity and reliability of data held against pre-defined criteria.5

DoCDat was accessed between May and August 2006 and a table detailing information on all datasets available online was compiled. Each dataset within the Directory was interrogated online by means of detailed searches of questionnaires and webpage information against pre-defined criteria using a standardized approach to ascertain whether or not ethnicity- and faith-related data were collected and, if so, which specific questions were used. If answers were unavailable through online searches, the custodian of the individual datasets was contacted by email in an attempt to obtain this information.

Data were abstracted onto a customized data extraction sheet and descriptive statistics were employed to summarize results.

Results
Online information was available for 95 of 162 datasets. Custodians of the remaining 67 datasets...
were emailed to request information on the variables of interest, and 37 responses were obtained. We were therefore able to ascertain whether questions on ethnicity and/or religion were included in 132 of 162 datasets (81%). Of the 132 datasets for which relevant information was available, 62 (46%) contained a question on ethnicity and seven (5%) on religion/fight.

The questions used to determine ethnicity varied widely, as demonstrated in Table 1. Of the 62 datasets collecting ethnicity data:

- 21 used an open question requesting details of ethnic group, ethnicity, ethnic code or ethnic category
- Nine used a question along the lines of the census question of 2001
- Five used a question with nine named categories

For the remaining 37 datasets it was not possible to ascertain the exact ethnicity question used.

All seven datasets collecting data on religion/fight used a different code, again reflecting a lack of standardization.

The email responses from data custodians revealed two major reasons for these data not being collected: perceived irrelevance of or difficulty in collecting such information; and a failure to appreciate the importance/potential uses of these data (Box 1).

Discussion

Despite ample evidence of persistent ethnic health inequalities, we have found that the majority of UK health datasets still fail to collect any data on
Box 1 Examples of email responses from data custodians

- Perceived lack of importance/difficulty ‘Too hard to get reliably to make it worthwhile collecting. And the South is still very white, although becoming less so...’ (R156)
- Failure to appreciate importance ‘Neither is included ... it would be very easy to add this information ...’ (R152) ‘... but now you have asked the question I will consider whether we will capture this in future’ (R42)

Health statistics fail to collect information on ethnicity and religion

In 2001, the UK Census and Home Office Citizenship Survey clearly demonstrated that it is feasible to collect valid and reliable data on ethnicity and religion, and that these data can be used for extremely important analyses. In, time on, it is time that the standard set in these national surveys is much more widely adopted within health-care services.

References


Box 1

Examples of email responses from data custodians

- Perceived lack of importance/difficulty ‘Too hard to get reliably to make it worthwhile collecting. And the South is still very white, although becoming less so...’ (R156)
- Failure to appreciate importance ‘Neither is included ... it would be very easy to add this information ...’ (R152) ‘... but now you have asked the question I will consider whether we will capture this in future’ (R42)