Research article

The collection and utilisation of patient ethnicity data in general practices and hospitals in the United Kingdom: a qualitative case study

Author address for correspondence:
Zoe Morrison
eHealth Research Group
Centre for Population Health Sciences
University of Edinburgh
Edinburgh EH8 9DX, UK
Email: zoe.morrison@ed.ac.uk

Zoe Morrison
eHealth Research Group, Centre for Population Health Sciences, University of Edinburgh, Edinburgh EH8 9DX, UK

Bernard Fernando
eHealth Research Group, Centre for Population Health Sciences, University of Edinburgh, Edinburgh EH8 9DX, UK

Dipak Kalra
Centre for Health Informatics and Multiprofessional Education, University College London, London, UK

Kathrin Cresswell
The School of Health in Social Science, University of Edinburgh, Edinburgh EH8 9DX, UK

Ann Robertson
eHealth Research Group, Centre for Population Health Sciences, University of Edinburgh, Edinburgh EH8 9DX, UK

Aziz Sheikh
eHealth Research Group, Centre for Population Health Sciences, University of Edinburgh, Edinburgh EH8 9DX, UK and Division of General internal Medicine and Primary Care, Brigham and Women's Hospital/Harvard Medical School, 3rd Floor, 1620 Tremont Street, Boston, MA 02120-1613, USA

ABSTRACT

Background Although the collection of patient ethnicity data is a requirement of publicly funded healthcare providers in the UK, recording of ethnicity is sub-optimal for reasons that remain poorly understood.

Aims and objectives We sought to understand enablers and barriers to the collection and utilisation of ethnicity data within electronic health records, how these practices have developed and what benefit this information provides to different stakeholder groups.

Methods We undertook an in-depth, qualitative case study drawing on interviews and documents obtained from participants working as academics, managers and administrators within the UK.

Results Information regarding patient ethnicity was collected and coded as administrative patient data, and/or in narrative form within clinical records. We identified disparities in the classification of ethnicity, approaches to coding and levels of completeness due to differing local, regional and national policies and processes. Most participants could not identify any clinical value of ethnicity information and many did not know if and when data were shared between services or used to support quality of care and research.

Conclusions Findings highlighted substantial variations in data classification, and practical challenges in data collection and usage that undermine the integrity of data collected. Future work needs to focus on explaining the uses of these data to front-line clinicians, identifying resources that can support busy professionals to collect standardised data and then, once collected, maximising the utility of these data.

Keywords: ethnicity, healthcare disparities, language, minority health, qualitative research, race

INTRODUCTION

Populations are becoming increasingly ethnically diverse and this is a trend that is set to continue. For example, UK data indicate that 14% of the total population in England and Wales categorise themselves as belonging to a minority ethnic group (Figure 1).[1] Many minority ethnic groups in Western countries have a different health profile to that of the general population,[2] and there is an increasing body of research demonstrating substantial and persistent variations in disease prevalence and outcomes by ethnic group, such as type 2 diabetes and coronary heart disease amongst South Asians,[3,4] sarcoidosis in people of African origin[5] and ethnic variations in asthma frequency and morbidity rates.[6,7] Routine capture and analysis of information on patient ethnicity have the potential to generate better data on these ethnic health variations and service provision to help improve delivery of equitable care. Ethnicity data are required to evidence compliance with anti-discrimination legislation,[8,9] for public health monitoring and for research into health inequalities,[3–7] including access to care for minority groups.[10,11] Clinical practice may also be informed by ethnicity, including the accommodation of cultural norms and community preferences that may impact upon healthcare delivery, such as a requirement for a same sex clinician.[12–14]

In some nations, such as France and Germany, the collection of data on ethnic origin is subject to restrictions.[15] In the United States, data on patient race, ethnicity and language were not previously routinely collected,[16] but more recently have been required as a result of state[17] and now federal legislation.[18] The UK officially recognises the importance of patient ethnicity data. The collection of demographic data relating to patient ethnicity in general practice has been encouraged within the Quality and Outcomes Framework (QOF)[19] (the national re-imbursement scheme for general practices in the UK (Box 1)) and has been a mandatory reporting requirement for hospitals since 1 April 1995.[20,21] Within clinical computer systems (i.e. electronic health records), data relating to patient ethnicity might be

---

Box 1 Definitions of key terms

Ethnicity

'The social group a person belongs to, and either identifies with or is identified with by others, as a result of a mix of cultural and other factors including one or more of language, diet, religion, ancestry, and physical features traditionally associated with race'.[41] Ethnicity is a highly subjective classification that an individual is required to articulate within a simple data item structure, and as such, it has been argued that the only true meaningful categorisation is self-definition.

Race

'The group (subspecies in traditional scientific usage) a person belongs to as a result of a mix of physical features such as skin colour and hair texture, which reflect ancestry and geographical origins'.[41] Race is now largely a discredited term in scientific circles as it fails to recognise the common genetic stock of people of different skin colours, emphasises above all physical characteristics, has been used to advance arguments in favour of biological determinism and undermines the socially constructed nature of many people's 'ethnic' identity.

Quality and Outcomes Framework (QOF)

This national framework pay for performance scheme was introduced in April 2004 as part of the new General Medical Services contract.[19]
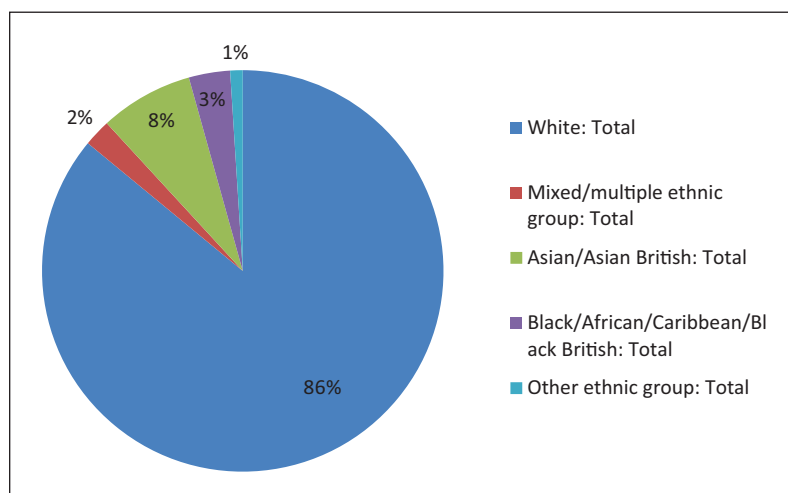

Figure 1 Summary of ethnic groups within England and Wales[1]

captured by healthcare professionals as part of the patient history as coded data (choosing from a pre-determined list with assigned numeric values for ease of computer retrieval and analysis) and/or documented in the clinical narrative as free text (and therefore not accessible to staff who need to report on ethnicity).

Despite government encouragement, the collection of ethnicity data remains sub-optimal in the UK.[22] Known issues include staff's lack of knowledge regarding the importance of the data,[23] (in)completeness driven by administrative processes[24] and (in)accuracy resulting from the use of overly simplistic unitary classifications.[25] In the context of a wider government-commissioned study on optimising the structuring/coding of digitised data in electronic health records, we undertook an in-depth evaluation of these issues in relation to ethnicity data. Our aim was to understand reasons for this impasse and how they might be overcome.

METHODS

Design

We conceptualised this research as a case study[26,27] of structuring and/or coding ethnicity data (Box 1). In view of the substantial body of work related to ethnicity data, we first reviewed existing literature and consulted extensively with relevant academic colleagues and reviewed relevant professional online discussions (MINORITY-ETHNIC-HEALTH within the UK National Academic Mailing List Service (https://www.jiscmail.ac.uk/cgi-bin/webadmin?A0=minority-ethnic-health)). We were also able to observe a two-hour multidisciplinary meeting of senior clinical and managerial staff with designated responsibility for the collection of ethnicity data within a regional health authority in Scotland. These preliminary activities informed our subsequent empirical investigation.

Ethics and governance

We obtained ethical approval for this work from the National Research Ethics Service – Brighton West Ethics Committee (MREC Ref: 10/H1111/25). Site-specific permissions from local National Health Service (NHS) research and development offices were facilitated by the Primary Care Research Network (PCRN) and the UK Clinical Research Network (UKCRN). Individuals expressing interest in participation were forwarded the study information sheet to allow them to further consider participation. Informed consent was gained in writing prior to face-to-face participant interviews, and verbal consent was recorded prior to commencement of telephone interviews.

Sampling and recruitment

We first invited three leading UK-based academics with interest in ethnicity data to participate in the study. We then worked with UKCRN and PCRN to purposefully sample[28] managers and administrators involved in processes of collection and utilisation of patient ethnicity data in general practices and hospitals and invited them to participate. Recruitment aimed to include those with expertise relating to ethnicity data and those with no particular expertise or previously declared interest in this area.

Data generation and handling

Our empirical data were derived from semi-structured interviews encompassing variations in profession, geographical locations and care settings, these including senior academic colleagues, senior managers, clinical coding managers, research and information professionals, general practice managers and administrators. We gathered documentation recommended to us by participants to gain an understanding of the range of current and recent work regarding patient ethnicity data.

Our interview topic guide was augmented and adapted during data generation to consider and clarify emergent themes (Table 1). Key issues explored included tools and techniques for the collection of ethnicity data, approaches to coding the information during data entry processes, external requirements and guidelines as barriers and enablers for data entry, perceptions regarding the relevance of ethnicity data and areas for possible improvement. Interviews were digitally recorded subject to participant consent. If participants chose not to be recorded, researcher field notes were taken. Recorded interviews were transcribed in full, anonymised and checked against the original recording. Data generation continued until saturation could reasonably be assumed.[29]

Analysis

Analysis was conducted iteratively using NVivo software.[30] During the period of data collection, the researcher (ZM) analysed the data generated using themes within the literature, then discussed findings with the research team, identifying areas for detailed investigation. Upon completion of data generation, further inductive analysis identified emergent themes, examples of particular interest and areas for clarification.[31] We actively sought disconfirming evidence, and adopted a reflexive approach to mitigate the influence of researcher prior knowledge and assumptions upon data analysis.[32]

RESULTS

We conducted 14 semi-structured interviews, 13 of which were conducted face to face and the remaining interview was conducted by telephone. Two of the face-to-face interviews were conducted with two interviewees, yielding a total of 16 participants. Participant characteristics are detailed in Table 2. We gathered 50 documents from participants (listed in Table 3). Findings are reported in relation to the collection and utilisation of patient ethnicity data. Where relevant, illustrative quotes from participants are given as transcribed from digital recordings. Where interviews were not audio-recorded, findings are presented based upon researcher field notes.

Table 1 Sample interview guide

Main structure	Specific topics and issues
Confidentiality, aims, thanks	Theorised and actual benefits and risks, drivers, incentives, barriers and how to address these.
Any questions?	
About yourself	Can you tell me a bit about your role and what you do here? (setting, profession, coding system, electronic system)
Understand local processes relating to the collection of patient ethnicity data.	Ask participants to describe how they collect and store patient ethnicity data, related documents, forms and local processes.
Do the structures and/or codes cover what you feel needs to be recorded – any areas for improvement?	In terms of completeness and accuracy and in terms of enabling good use of the information.
	Any potential uses of the information that are under-exploited? If yes, why?
Overall	
How well do the available clinical systems support structuring and/or encoding the information?	
Any other barriers to collecting good quality information?	
Any drivers or incentives that would improve the quality or uses made of this information?	
Any developments in relation to structuring and/or coding patient ethnicity data they are aware of?	Any examples of innovation/centres of excellence?
	Aware of particular practical issues or areas of concern?
Concluding remarks	
Anything else?	Anyone they can recommend for interview?
	Any relevant literature?
Thanks, any questions or further things you would like to discuss?	

Table 2 Interviewee participant characteristics

Identifier	Role	Gender	Setting	Interview method
1	Senior Academic	Male	Research Institute	Face-to-face interview
2	Senior Academic	Male	Research Institute	Face-to-face interview
3	Senior Academic	Male	Research Institute	Face-to-face interview
4	Senior Manager	Female	NHS Scotland	Face-to-face interview
5	Clinical Coding Manager	Male	Hospital	Face-to-face interview
6	Research Nurse	Female	Hospital	Face-to-face interview
7	Information Manager	Male	Hospital	Face-to-face interview
8	Senior Manager	Male	NHS England	Face-to-face interview
9	Practice Manager	Female	General practice	Face-to-face interview
10	Administrator	Female	General practice	Face-to-face interview
11	Practice Manager	Female	General practice	Face-to-face interview
12	Administrator	Female	General practice	Face-to-face interview
13	Practice Manager	Female	General practice	Face-to-face interview
14	Information Manager	Female	Hospital	Face-to-face interview
15	Senior Manager	Female	Hospital	Face-to-face interview
16	Clinical Coding Manager	Female	Hospital	Telephone interview

Table 3 Documents collected from participants

1.	List to reflect populations in East London (Template for ethnic category – 2011 census)
2.	EMIS national template (Template for Template GMS – Ethnicity V12)
3.	Version 4 – Jan 2008 GP Contract and Enhanced Service Ethnicity – Patient Profiling Template Guide
4.	Patient Profiling Form
5.	[General Practice] HC Audit Results Enhanced Services 2009 and 2010
6.	Health Equality in primary care: Cardiovascular disease, diabetes and COPD in inner east London
7.	[General practice] chronic disease management 2004–2009
8.	Clinical Coding Clinicians Nov 2010
9.	[Management group] Communications Plan Aug 11 update ADJG
10.	Final EDIP Update July 2011
11.	[Management Group] primary care examples of clinical relevance
12.	[Management Group] Checklist primary care Jun 11
13.	[Management Group] Agenda 27 Sep 2011
14.	[Management Group] Action Notes 290611 final
15.	[General practice] report Sep 10
16.	SMR Ethnicity codes and date ranges
17.	Ethnicity [region] Hospital Figures 190911–250911
18.	Edinburgh Ethnicity and Health Research Group Meeting 060911
19.	The Essential Guide to the Public Sector Equality Duty
20.	Culturally responsive JSNAs: a review of race equality and Joint Strategic Needs Assessment (JSNA) practice
21.	Collecting Ethnic Category Data
22.	Fairness and Freedom: The Final Report of the Equalities Review
23.	Department of Health Data Standards: Ethnic Category
24.	Ethnic Category Standard (v3.0.1) Change to an Information Standard
25.	Ethnic Category Standard (v3.0.1) Change to an Information Standard Human Behavioural Guidance
26.	Relevant Info from Information Standards Board for Health and Social Care
27.	Practical Guide to ethnic monitoring in the NHS and social care
28.	Summary of Responses to consultations
29.	Improving data collection for equality and diversity monitoring all Scotland
30.	Collecting Equality Information Guidance on asking questions on: Ethnic Groups
31.	Equality and diversity
32.	Scotland's New Official Ethnicity Classification
33.	Meeting notes from [hospital name] 270111
34.	Problems and Barriers to the collection of ethnicity and migration related data: the UK experience
35.	HES Data Dictionary
36.	Recording of Ethnic Group: Information for Patients [hospital name]
37.	Date about ethnicity: Information for patients and carers [hospital name]
38.	GP Contract and Enhanced Service Ethnicity – Patient Profiling Template Guide
39.	Data Entry Codes available on EMIS/Vision
40.	New Patient Registration Form [general practice]
41.	New Patient Registration Form [general practice]
42.	Medway DES – Ethnicity [general practice]
43.	Data Quality Audit Ethnicity Apr 2011 [general practice]
44.	Ethnicity Protocol [general practice]
45.	New Patient Registration Form [general practice]
46.	Data Quality Audit Ethnicity 2011 [general practice]
47.	Patient Participation Questionnaire [general practice]
48.	[Region] – English not First Language Service Level Agreement
49.	Taskforce Summary 270911
50.	Ethnicity recorded [management group] report Sep 2011

Collection of patient ethnicity data

Classification of ethnicity data

Patient ethnicity data were asked for and recorded either by administrative staff as part of the patient demographic data used for administrative purposes and/or by health care professionals as information gathered within the clinical encounter. Administrative staff usually collected demographic data during new patient registration in general practices and upon first appointment or admission in hospitals. These data were often stored separately from clinical systems, and so were not always accessible to clinicians. Demographic data were almost always requested of patients using a paper form as part of either practice registration or first attendance within an episode of hospital care. A single ethnic category was either chosen by patients from a pre-determined list or self-defined in response to an open question (see Table 4, for examples of data collection forms from general practices). The response given was then input into computer records as one single data item: ethnic origin. Patients were given the option to decline to state their ethnicity. In both care settings, ethnicity data collection were seen as an opportunity to ask patients about any need for interpreter services and this information was grouped within the same data entry screen as ethnicity.

We noted variations in the classification of ethnicity data. Ethnicity classifications were derived from the UK's national decennial censuses.[25] Surprisingly, ethnicity classifications were not the same across the four home nations (i.e. England, Scotland, Wales and Northern Ireland) comprising the UK. Whilst new categories were introduced into Scotland from 1st April 2012, categories in use in England remained the same. Variations were also evident between general practices and hospitals. In general practices, electronic health records were populated using Read Codes (from the clinical coding scheme currently used in UK general practice computerised records).[25] We identified some 83 codes available for data entry, in addition to the code accommodating patients who chose not to state their ethnicity. These codes are reproduced in Table 5. In contrast, recording in hospitals was undertaken using nationally determined data sets derived from census categories. In hospitals in England, the '16+1' categories were in use, with the 'one' representing the patient's prerogative to choose not to answer this question[33] whilst in Scottish hospitals '19+1' categories were in use.[34] These codes are reproduced in Table 6.

Completeness of demographic data

Levels of completeness of ethnicity data and approaches to coding were subject to variation due to differing local, regional and national policies and processes. Local factors such as individual general practice responses to financial incentive schemes, patient turnover rates and local collection mechanisms, staff training and support all contributed to variable data quality.

Incentivising the recording of patient ethnicity data for patients in general practices were previously included within the QOF national payment scheme (Box 1), but this ceased in 2011 (Interview 8, Senior Manager, NHS England). Some participants found this omission frustrating, particularly as it led to its removal as an option within the standard electronic health record reporting menu for local audit of data completeness:

'But it's a bit annoying that I can't, I can't see, they've taken the audit off'. (Interview 13, Practice Manager)

A possible reason for this change was that the collection of ethnicity data was felt to have become embedded in normal working practice, and therefore the incentive was no longer necessary. We explored this amendment to the payment incentive framework (i.e. QOF) with an interviewee from the Department of Health in England, who explained the purpose of the quality drivers:

'QOF is to drive behaviour not sustain it'. (Interview 8, Senior Manager, NHS England)

Some local initiatives also sought to drive improvements in completeness of data collection. These schemes were founded on the principle that collecting patient ethnicity data are simply *'The right thing to do'*. (Interview 2, Senior Academic). These schemes were felt to be positively influencing practice:

'Scotland had delivering enhanced services schemes that practices could choose to do for extra income. Ethnicity was one of them for total patient population. They were aiming for certain percentages, between 80% and 90% [coverage]. Not all practices chose to do this and [it is] difficult to know how many did. Estimated 69% of practices took this up [approximately 1,030 practices in Scotland]'. (Interview 4, Senior Manager, NHS Scotland).

All hospital-based participants knew of the requirement to report information on a monthly basis as part of the main contractual data flows to the Department of Health Information Centre for secondary uses. Interviewees were less aware of the need for data completeness and knew of no sanction for failing to collect these data (Interviews 7 and 14, Hospital Information Managers). The amount of information coded varied depending on the speciality, suggesting higher rates of completion in some clinical specialties (e.g. when documenting ethnicity for a disease register). This may have been due to enhanced expectations of high levels of data completeness, perhaps supported by the design of speciality specific computer applications (Interview 1, Senior Academic). Thus, an individual hospital's overall recording levels may have been misleading as they were based on an average figure for the organisation that may mask differences in levels of completeness across different services within a hospital.

Accuracy of demographic data

Templates used in hospitals for collecting data on self-defined ethnicity were devised locally, and were similar due to the mandating of data collection and the restricted number of codes used. In general practices, we found no consistent format for the collection of patient ethnicity information from patients and practices varied greatly. In England, there was

Table 4 Examples of general practice patient ethnicity data collection forms

New patient questionnaire	Patient ethnic origin questionnaire	New patient questionnaire	Patient profiling form	
WHITE	**White**	Ethnicity	White British	24 Language options listed
British	British		White Irish	
		Main first spoken language		**If your service is provided in English do you need an interpreter/advocate? (Yes or No)**
Any other white background – please specify	Irish		Traveller	Albanian
	Any other White background please write in below		Greek	Bosnian
ASIAN OR BRITISH ASIAN			Turkish	Croatian	**If you have a visual impairment do you require?**
Indian/British Indian	**Mixed**		Kurdish	Serbian	Braille (Yes/No)
Pakistani/British Pakistani	White and Black Caribbean		Estonian/Latvian/ Lithuanian	Other Yugoslav	Large Print (Yes/No)
Bangladeshi/British Bangladeshi	White and Black African		Russian	Other White	
Any other Asian background – please specify	White and Asian				**What is your faith or religion, if any?**
	Any other Mixed background please write in below				14 Options, including Agnostic, None and Religion declined
MIXED			Jewish	British Asian	
White and Black Caribbean	**Asian or British Asian**		Bengali/British Bengali	Mixed Asian	
White and Black African	Indian		Pakistani/British Pakistani	Chinese	
White and Asian	Pakistani		Indian/British Indian	Vietnamese	
Any other Mixed background – please specify	Bangladeshi		Sri Lankan	Japanese	
	Any other Asian background please write in below		Tamil	Filipino	
BLACK OR BLACK BRITISH			Sinhalese	Other Asian	
Black Caribbean/British Caribbean	**Black or Black British**		Black African	Mixed Black	
Black African/British African	Caribbean		Black Caribbean	Mixed White/Black Caribbean	
Any other Black background – please specify	African		Black British	Mixed White/Black African	
OTHER ETHNIC GROUP	White and Asian		Somali	Mixed White/Asian	
	Any other Black background please write below		East African Asian	Any other Mixed	
Chinese	**Chinese or other ethnic group**		Caribbean Asian	Other non-mixed	
Other – please specify	Chinese or other ethnic group		Middle Eastern		
I DO NOT WISH TO ANSWER	Any other please write below			I do not wish to state my ethnicity	
Do you need an interpreter?	First Language		Other Black		
Please state your first language					

Table 5 **Example of general practice ethnicity data computer entry options in England**

Ethnicity not given – patient refused	Other mixed background	Other ethnic category
British or Mixed British	Other Mixed background	Chinese and White
British or Mixed British	Black and Asian	Vietnamese
Irish	Black and Chinese	Japanese
White or Mixed White	Black and White	Filipino
Other White background	Chinese and White	Malaysian
English	Asian and Chinese	Buddhist
Scottish	Other Mixed or Mixed unspecified	Hindu
Welsh	White and Black Caribbean	Jewish
Cornish	White and Black African	Muslim
Northern Irish	White and Asian	Sikh
Ulster Scots	**Asian or Mixed Asian**	Arab
Cypriot (part not stated)	Indian or British Indian	North African
Greek	Pakistani or British Pakistani	Mid East (excl. Israeli, Iranian and Arab)
Greek Cypriot	Bangladeshi or British Bangladeshi	Israeli
Turkish	Other Asian Background	Iranian
Turkish Cypriot	Punjabi	Kurdish
Italian	Kashmiri	Moroccan
Irish Traveller	East African Asian	Latin American
Traveller	Sri Lankan	South and Central American
Gypsy/Romany	Tamil	Mauritian/Seychellois/Maldivian/St. Helena
Polish	Sinhalese	
Baltic Estonian/Latvian/Lithuanian	Caribbean Asian	**Language**
Commonwealth (Russian) Independent States	British Asian	Additional main spoken language
Kosovar	Mixed Asian	English as a second language
Albanian	Other Asian or Asian unspecified	[click for list of languages]
Bosnian	**Black or Mixed Black**	**Interpreter required?**
Croatian	Caribbean	[click for list of languages]
Serbian	African	
Other Republics former Yugoslavia	Other Black background	
Mixed Irish and Other White	Somali	
Other White European/European unspecified/Mixed European	Nigerian	
Other Mixed White	Black British	
Other White or White unspecified	Mixed Black	
	Other Black or Black Unspecified	

Table 6 Mandated patient ethnicity data codes in secondary care in England and Scotland

NHS England ethnic category national codes[33]	NHS Scotland ethnicity codes in secondary care[34]
White	**Group A – White**
A British	1A Scottish
B Irish	1B Other British
C Any other White background	1C Irish
Mixed	1K Gypsy/Traveller
D White and Black Caribbean	1L Polish
E White and Black African	1Z Other White ethnic group
F White and Asian	**Group B – Mixed or multiple ethnic groups**
G Any other Mixed background	2A Any Mixed or multiple ethnic groups
Asian or Asian British	**Group C – Asian, Asian Scottish or Asian British**
H Indian	3F Pakistani, Pakistani Scottish or Pakistani British
J Pakistani	3G Indian, Indian Scottish or Indian British
K Bangladeshi	3H Bangladeshi, Bangladeshi Scottish or Bangladeshi British
L Any other Asian background	3J Chinese, Chinese Scottish or Chinese British
Black or Black British	3Z Other Asian, Asian Scottish or Asian British
M Caribbean	**Group D – African**
N African	4D African, African Scottish or African British
P Any other Black background	4Y Other African
Other ethnic Groups	**Group E – Caribbean or Black**
R Chinese	5C Caribbean, Caribbean Scottish or Caribbean British
S Any other ethnic group	5D Black, Black Scottish or Black British
Other codes	5Y Other Caribbean or Black
Z Not stated	**Group F – Other ethnic group**
99 Not known	6A Arab, Arab Scottish or Arab British
	6Z Other ethnic group
	Group G – Refused/Not provided by patient
	98 Refused/Not provided by patient
	Group H – Not Known
	99 Not Known

a lack of useful resources and we found no widely known support to facilitate consistent and accurate data collection and entry by general practices. We observed that collection of patient ethnicity data was not supported by the Family Doctor Services Registration Form (GMS1) used in new patient registration when the patient did not have a medical card, as is often the case for vulnerable members of minority groups, such as asylum seekers and immigrants. We collected four registration forms designed locally to capture the relevant information on paper, before data entry (Table 4), illustrating the degree of variation between practices. When the data entry options detailed in Table 5 are compared with options listed on registration forms (Table 4), we see the extent of interpretation necessary, and potential for inconsistency, in mapping these categories to those available for data entry.

Other challenges to consistent data collection were felt to lie in obtaining the information from patients due to staff embarrassment, fear of causing offense and in some cases perceived irrelevance (particularly in the context of emergency care provision by, for example paramedical staff). One suggestion for obtaining the information from patients was to change the point in the care pathway, at which information is collected from the beginning to the end:

'*In hospital it may be easier to ask in hospital discharge – because you're ticking sheets anyway*'. (Interview 4, Senior Manager, NHS Scotland)

Utilisation of patient ethnicity data
Clinical relevance

In clinical contexts, academic participants suggested potential for wider, socially patterned ethnicity-related data relating more closely to the intended use of the information, including information on country of birth, languages spoken, religious affiliation, diet, nationality, citizenship and migration status (Interviews 1–3, Senior Academics). This information was considered potentially important to the provision of personalised clinical care and, if overlooked, could result in sub-optimal care. However, we were unable to locate a recommended format for the consistent, structured recording of these data by clinicians. These items were described as being collected at the clinician's discretion within the clinical narrative and recorded in free text formats in electronic health records that were not shared between care settings.

Many interviewees working in general practices and hospitals did not know if or how patient ethnicity information was shared between services or used in the provision of care:

'The demographics is just a surname, forename, gender, date of birth and that's how we would find a patient but again there's no, nothing on there for ethnic origin. I don't know whether they ask that at the hospital when they attend'. (Interview 13, Practice Manager)

'I can't think that we do anything to the hospital, I think they do their own monitoring but we certainly don't provide it or we're not asked to provide it to the hospital … All they know is how many people we've recorded but they've never asked me for a breakdown of the actual groups'. (Interview 7, Information Manager)

'The only other thing is for diabetics who are, I suppose that's more religion I suppose, obviously being diabetic and fasting'. (Interview 12, Practice Administrator)

Despite governmental guidelines and legislation, such as *The Essential Guide to the Public Sector Equality Duty* [35] and the Equality Act 2010, we could not identify any national agreements, recommendations, or resources to support processes for general practices to communicate information on patient ethnicity in referrals to hospital or other care providers. Conversely, we identified local examples of good practice in England and Scotland (Box 2).

Box 2: Examples of good practice

London borough

One borough in London had worked hard to capture patient ethnicity information which they saw as fundamental to their work given the rich diversity of their community, including the development of templates, interpreter services and staff training.

'For the last almost 10 years now we have been promoting the recording of self-reported ethnicity by patients in GP [general practice] records and we've set up coding structures, templates that allow that for all new patient registrations, six week baby checks, chronic disease management have all got these templates built in, staff have been trained in practices and it's been supported by…you know financially by the enhanced service, by the three PCTs [Primary Care Trusts – groups of general practices] such that we've now got over 80% of self-reported ethnicity recording in the total population and over 90% in people with long term conditions so we have complete ethnicity recording across 800,000 people in the most disadvantaged, socially diverse population in the UK'. (Interview 1, Senior Academic)

Scotland

Previous analyses of recording levels across Health Boards in Scotland showed extreme variations and shortcomings in levels of completeness of patient ethnicity information recording. These issues in data collection were addressed within a national programme, developed and implemented over a 10-year period, comprising a six strand national initiative to increase levels of coding and included the development of an Ethnic Monitoring Toolkit, including training resources and handbooks for staff:

'We had six work streams:
- *originally about getting IT in place;*
- *classifications needed to be right;*
- *training resources;*
- *PFPI(Patient Focus and Public Involvement) engagement – BME (Black and Minority Ethnic) communities, learning disabilities, citizens panel;*
- *research group;*
- *human resource'.*
  (Interview 4, Senior Manager, NHS Scotland)

Local initiatives were working to support the national programme, and we observed the meeting of a multi-disciplinary working group convened by a Health Board to increase the levels of completeness of patient ethnicity data collection. Their objective was to increase ethnicity coding in general practices and hospitals to 90% completeness within three years. This represented a shift from a starting point of 5% completeness in hospitals and 35% completeness in general practices. Whilst the group was making a significant difference in hospitals due to local policy directives, systems changes and clinical leadership, they were less able to influence the more autonomous general practitioners.

Patient access and participation

One area where ethnicity data were seen as useful was in making services more accessible to patients. This was felt to be particularly relevant for patients with language considerations:

'It's just a data exercise really and as I say we're just noticing ethnic minorities coming into the practice so I'm having to think ahead, I'm thinking right what changes am I going to have to make, i.e. material that we have in our waiting room, making sure that I know how to get hold of an interpreter for, it's just things like that that I'm having to be aware of now'. (Interview 11, Practice Manager)

This was an unanticipated finding as ethnicity *per se* is a poor indicator of the need for an interpreter given the presence of many second, third or more generation immigrants and the increasingly diverse nature of the population.

Participants working within general practice articulated concerns regarding patient non-participation in preventive health care initiatives due principally to language and/or cultural considerations. Examples cited included: very low rates of attendance for routine cervical screening by women which was attributed to patients being unable to read the letters of invitation; failure of children to attend for immunisation due to similar problems with communication; non-participation in breast screening programmes by some women due to the involvement of male staff; and similar barriers to participation in faecal occult blood testing for bowel cancer screening (Interview 9, Practice Manager; Interview 10, Administrator). When asked if they had used ethnicity data available within the practice to verify these concerns, interviewees said they had been unable to do this due to lack of available administrative staff time and resources to support local initiatives, such as the translation of practice leaflets.

Participants with a particular interest in health inequalities were concerned by the potential risks arising from a lack of data or data that were increasingly out of date and unavailable to service commissioners (Interview 2, Senior Academic). A more general risk noted was the potential for legal challenge to service providers based on alleged contravention of equalities legislation (for example Documents 19 and 29), although no interviewee recalled an example of an action being taken against a UK-based health care provider.

Uses for secondary analysis

We identified a number of possible secondary uses of ethnicity data, particularly in the context of epidemiological research (Interviews 1–3, Senior Academics), such as population tracking and disease prevalence amongst sub-populations. These uses were described by researchers as highly cost-effective contributions to the development of policy and practice in the UK, particularly given the likely cessation of the decennial census. Collection of these data were felt to be critical to the reduction of health inequalities (Interviews 1–3, Senior Academics), offering significant potential for cost-effective advances in knowledge utilising data linkage techniques to

further investigate phenomenon of interest, such as patterns in disease incidence and prevalence; access to, and the efficacy of, health care services and interventions. Ironically, those we spoke to in hospitals and general practices working to capture this valuable information seemed largely unaware of their contribution to this paper.

DISCUSSION

Summary of main findings

Our findings highlighted notable inconsistency between the patient ethnicity coding schemes used in hospitals and general practice, and between classifications used in hospitals in England and Scotland. We also identified variations in local arrangements for the collection of patient ethnicity data in general practices and, to a lesser degree, hospitals. This was due in part to contextual factors, including lack of training for staff, resource constraints, variations in the extent of ethnic diversity in different regions and local working practices. We noted a lack of central provision of supporting resources, such as templates for data collection, training and development incentives and/or sanctions for organisations collecting this information. The potential benefits and usefulness of this data for secondary analysis were described as significant, although lack of completeness of data and inconsistencies in classification were felt to limit opportunities for research. Our findings indicate that the benefits of collecting and using patient ethnicity data are currently limited by a lack of policy emphasis, an absence of data sharing across care settings, a perceived lack of organisational and clinical relevance, and pronounced variations in emphasis on equality and diversity and knowledge of data usage.

Strengths and limitations

This case study gathered perspectives from academic experts, clinicians, managerial and administrative staff to consider the collection of ethnicity data. We have sought to extend previous work[15,20,22,36,37] by considering in detail examples of local arrangements for the collection and use of ethnicity data in the context of relevant national requirements and guidelines. We have also considered perceptions as to the clinical relevance and use of these data, and the potential for more meaningful construction of ethnicity related information to enable personalised clinical care. Although a UK case study, advances made by the NHS suggest recommendations are likely to be relevant across a range of health care services, where electronic health records are being used and ethnicity monitoring is mandated and/or incentivised, such as in the US.[17,18] The main limitations of this study were our small sample size and focus on demographic data relating to patient ethnicity. We did not examine medical records to consider the capture of data relating to ethnicity within clinical narratives, nor interview patients and non-expert (i.e. without a special interest in health inequalities) clinical staff regarding the accommodation of ethnicity related factors within the provision of personalised care. This paper is, so far as we can establish, yet to be done and we recommend it as an important area for further investigation.

Exploring our findings in the light of the existing literature

Our findings confirm existing work in relation to the completeness of ethnicity data within patient records in the UK and elsewhere.[15,20,22] We found significant variations in classification and collection of demographic ethnicity data, confirming known practical difficulties.[36,37] These variations may have been masked by discussions regarding completeness of ethnicity data as opposed to consistency, i.e. a focus on ensuring a data entry for each patient without due consideration of processes of interpretation and abstraction during collection and electronic data entry. That said, we noted potential value to be gained from variations in data collection regarding ethnicity to reflect the needs of local communities. For example, the requirements for palliative care[11] may be very different from the information needed by paramedics, yet our findings indicate that currently these data are determined according to the same mandated national reporting requirements in England[33] and Scotland.[34]

Implications for policy, practice and research

Financial incentives in general practice and mandating of data collection in hospitals have driven nationwide implementation of the collection of ethnicity data in the UK.[15,20,36,37] This research has uncovered a number of drivers and benefits from the use of information that relates to patient ethnicity together with somewhat haphazard attitudes and working practices. The primary legislative driver for ethnicity recording is to enable demonstration of equity of access to health care services across ethnic groups, and equality of care provided. In societies in which there are multiple generations of individuals who will have the same ethnic group but very different capabilities of handling the health service, the main challenge is not to ensure equity of access across ethnic groups, but to ensure that all persons have equal access to services regardless of any difficulty, they might have with making such access. We share concerns that reliance upon a single data item resulting from inconsistent processes is not an effective way of meeting this need.[21,25] We recommend a new impetus be given to the collection of ethnicity data to capitalise on the accomplishments achieved in the NHS and build capacity for the planning and commissioning of services responsive to the individual care needs and health promotion priorities of the population as it changes over time. This could be achieved by a combination of national support, for example the provision of supporting resources, such as data collection templates and local initiatives, such as training. Arguments for the collection of such data should not, however be conflated with clinical considerations – these demographic data need to be collected, wherever possible by non-clinical staff or perhaps even by self-completion by patients, for example using self-service kiosks.

Somewhat surprisingly, despite compelling arguments for the clinical relevance of data relating to patient ethnicity in the provision of personalised care[4,12,13,38,39] (e.g. to stratify patients biologically (genetically) into risk groups for certain diseases, optimise treatment plans; provide culturally sensitive health care and lifestyle guidance, and to provide health care in appropriate languages for those who cannot engage in a clinical consultation in English), how and when this information is captured within clinical narratives has not been studied for completeness and effectiveness in supporting clinical decision-making. This is at odds with guidance recommending the collection of this data as good practice, for example that of professional bodies, such as the British Medical Association.[40] Of course, the items in this list are not all resolved through a single ethnic group code: a person's racial group, cultural ethnicity, religion, spoken English language proficiency and written English language proficiency may all have different answers. Whilst these are not established methods of ethnic monitoring, we suggest that high-quality data of this nature would be a far better way of demonstrating equity of access and equal quality of care than simple tagging of ethnic group. We recommend this as a priority area for further investigation to inform the development of patient-centred approaches to care and the reduction of health inequalities.

CONCLUSION

Our findings give insight into the practical difficulties relating to the collection of patient ethnicity data together with variation and inconsistency in classification and consequent utility of the data collected. Although the collection of ethnicity data are recognised by some as important for legal, public health and academic purposes, realising the full potential of using ethnicity-related information for clinical care is probably limited given current variations in practice. A common set of categories in use across a health care community is needed to facilitate both consistency of data and efficient data sharing. Findings have also identified a range of wider ethnicity-related information regarding, for example the ability to speak or write in English and dietary requirements which may impact more directly on quality of care. Structured collection and codification of wider, ethnicity-related data, rather than free text capture, has potential clinical value, and consideration should therefore be given to better supporting the classification, collection, capture and use of such data.

Competing interests
None declared.

REFERENCES

1. Official Labour Market Statistics. 2011 Census Analysis. London. URL: http://www.nomisweb.co.uk/census/2011/lc2101ew [last accessed 01/08/13].

2. Office of National Statistics. Ethnic Variations in General Health and Unpaid Care Provision. London. URL: http://www.ons.gov.uk/ons/dcp171776_318773.pdf [last accessed 01/08/13].

3. Fischbacher CM, Bhopal R, Steiner M, Morris AD and Chalmers J. Is there equity of service delivery and intermediate outcomes in South Asians with type 2 diabetes? Analysis of DARTS database and summary of UK publications. *Journal of Public Health* 2009;31(2):239–49. doi: 10.1093/pubmed/fdp003. http://dx.doi.org/10.1093/pubmed/fdp003.

4. Fischbacher CM, Bhopal R, Povey C, Steiner M, Chalmers J, Mueller G et al. Record linked retrospective cohort study of 4.6 million people exploring ethnic variations in disease: mycocardial infarction in South Asians. *BMC Public Health* 2007;7:142. http://dx.doi.org/10.1186/1471-2458-7-142. PMid:17615055; PMCid:PMC1965474.

5. Sims HS and Thakkar KH. Airway involvement and obstruction from granulomas in African–American patients with sarcoidosis. *Respiratory Medicine* 2007;101(11):2279.

6. Sheikh A. Recruitment of ethnic minorities to asthma studies. *Thorax* 2004;59:634. PMid:15223878; PMCid:PMC1747077.

7. Sheikh A, Halani L, Bhopal R, Netuveli G, Partridge MR, Car J et al. Facilitating the recruitment of minority ethnic people into research: qualitative case study of South Asians and asthma. *PLoS Medicine* 2009;6(10):e1000148. doi: 10.1371/journal.pmed. http://dx.doi.org/10.1371/journal.pmed.1000148.

8. Williams C and Johnson MRD. Race and Ethnicity in a Welfare Society. Maidenhead: Open University Press, 2010.

9. Bhopal R, Fischbacher CM, Povey C, Chalmers J, Mueller G, Steiner M et al. Cohort profile: Scottish health and ethnicity linkage study of 4.65 million people exploring ethnic variations in disease in Scotland. *International Journal of Epidemiology* 2010;1–8.

10. Thorlby R, Jorgensen BS and Ayanian JZ. How health care organisations are using data on patients' race and ethnicity to improve quality of care. *Millbank Quarterly* 2011;89(2):226–55. http://dx.doi.org/10.1111/j.1468-0009.2011.00627.x. PMid:21676022; PMCid:PMC3142338.

11. Worth A, Irshad T, Bhopal R, Brown D, Lawton J, Grant E et al. Vulnerability and access to care for South Asian Sikh and Muslim patients with life limiting illness in Scotland: prospective longitudinal qualitative study. *British Medical Journal (Clinical research ed)* 2009;338:b183. doi:10.1136/bmj.b183. http://dx.doi.org/10.1136/bmj.b183.

12. Partridge MR. In what way may race, ethnicity or culture influence asthma outcomes? *Thorax* 2000;55:175–6. http://dx.doi.org/10.1136/thorax.55.3.175. PMid:10679533; PMCid:PMC1745701.

13. Netuveli G, Hurwitz B, Levy M, Fletcher M, Barnes G, Durham SR et al. Ethnic variations in UK asthma frequency, morbidity, and health-service use: a systematic review and meta analysis. *Lancet* 2005;365:312–7. http://dx.doi.org/10.1016/S0140-6736(05)70195-1. http://dx.doi.org/10.1016/S0140-6736(05)17785-X.

14. Lawton J, Ahmad N, Hanna L, Douglas M and Hallowell N. 'I can't do any serious exercise': barriers to physical activity amongst people of Pakistani and Indian origin with Type 2 diabetes. *Health Education Research* 2006;21(1):43–54. http://dx.doi.org/10.1093/her/cyh042. PMid:15955792.

15. Raleigh VS. Collection of data on ethnic origin in England is improving, but information needs to be acted on for health inequalities to narrow. *British Medical Journal (Clinical research ed)* 2008;337:a1107. doi: 10.36/bmj.a. http://dx.doi.org/10.1136/bmj.a1107.

16. Fiscella K, Franks P, Gold M and Clancy CM. Inequality in quality: addressing socioeconomic, racial and ethnic disparities in health care. *The Journal of the American Medical Association* 2000;283:2579–84. http://dx.doi.org/10.1001/jama.283.19.2579

17. Jorgensen S, Thorlby R, Weinick RM and Ayanian JZ. Responses of Massachusetts hospitals to a state mandate to colelct race, ethnicity and language data from patients: a qualitative study. *BMC Health Services Research* 2010;10:352. http://dx.doi.org/10.1186/1472-6963-10-352. PMid:21194450; PMCid:PMC3022878.

18. Ulmer C, McFadden B, Nerenz DR (Eds.) Race, ethnicity and langauges data: standardization for health care quality improvement. Institute of Medicine Report. 2010. URL: http://www.ahrq.gov/research/findings/final-reports/iomracereport/index.html [last accesssed 30/07/2013].

19. Boeckxstaens P, De Smedt D, De Maeseneer J, Annemans L and Willems S. The equity dimension in evaluations of the quality and outcomes framework: A systematic review. *BMC Health Services Research* 2011;11:209. http://dx.doi.org/10.1186/1472-6963-11-209.

20. Aspinall PJ. The mandatory collection of data on ethnic group of inpatients: experience of NHS trusts in England in the first reporting years. *Public Health* 2000;114(4):254–9. http://dx.doi.org/10.1038/sj.ph.1900662. http://dx.doi.org/10.1016/S0033-3506(00)00340-1.

21. Psoinos M, Hatzidimitriadou E, Butler C and Barn R. Ethnic Monitoring in Healthcare Services in the UK as a Mechaism to Address Health Disparities: A Narrative Review. London: Swan IPI, 2011. ISBN:978-0-9558329-8-7.

22. Sultana K and Sheikh A. Most UK datasets of routinely collected health statistics fail to collect information on ethnicity and relegion. *Journal of the Royal Society of Medicine* 2008;101:463–5. http://dx.doi.org/10.1258/jrsm.2008.080007.   PMid:18779248; PMCid:PMC2587383.

23. Iqbal G, Gumber A, Szczepura A, Johnson MRD, Wilson S and Dunna JA. Improving ethnic data collection for statistics of cancer incidence, management, mortality and survival in the UK. Final Report. London: Cancer Research UK, 2008.

24. Saunders CL, Abel GA, Turabi AE, Ahmed F and Lyratzopoulos G. Accuracy of routinely recorded ethnic group information compared with self-reported ethnicity: evidence from the english cancer patient experience survey. *British Medical Journal Open* 2013;3:e002882. doi: 10.1136/bmjopen-2013-002882. http://dx.doi.org/10.1136/bmjopen-2013-002882.

25. Aspinall PJ. The operationalization of race and ethnicity concepts in medical classification systsms: issues of validity and utility. *Health Informatics Journal* 2005;11:259–74. http://dx.doi.org/10.1177/1460458205055688.

26. Crowe S, Cresswell K, Robertson A, Huby G, Avery A and Sheikh A. The case study approach. *BMC Medical Research Methodology* 2011;11(1):100.   http://dx.doi.org/10.1186/1471-2288-11-100. PMid:21707982; PMCid:PMC3141799.

27. Stake RE. The Art of Case Study Research. London: Sage Publications,1995.

28. Silverman D. Doing Qualitative Research: A Practical Handbook. London: Sage, 2000.

29. Morse J. The significance of saturation. *Qualitative Health Research* 1995;5(2):147–9. http://dx.doi.org/10.1177/104973239500500201.

30. QSR International Pty Ltd. *NVivo* qualitative data analysis software Version 9. 2008.

31. Mason J. Qualitative Researching. London: Sage, 2002.

32. Alvesson M and Skoldberg K. Reflexive Methodology: New Vistas for Qualitative Research. London: Sage, 2000.

33. Department of Health Information Standards Board for Health and Social Care Data Standards: Ethnic Category. URL: http://www.isb.nhs.uk/documents/dscn/dscn2008/dataset/112008.pdf [last accessed 30/07/2013].

34. NHS National Services Scotland. Information Services Division Data Dictionary. URL: http://www.datadictionaryadmin.scot.nhs.uk/Dictionary-A-Z/Definitions/index.asp?Search=E&ID=243&Title=Ethnicity%20Code [last accessed 30/07/2013].

35. Equality and Human Rights Commission. The Essential Guide to the Public Sector Equality Duty: England (and non-devolved public authorities in Scotland and Wales). URL: http://www.equalityhumanrights.com/uploaded_files/EqualityAct/PSED/essential_guide_update.pdf [last accessed 19/02/14].

36. Pringle M and Rothera I. Practicality of recording patient ethnicity in general practice: descriptive intervention study and attitude survey. *British Medical Journal (Clinical research ed)* 1996;312:1080–2.   http://dx.doi.org/10.1136/bmj.312.7038.1080. PMid:8616418. PMCid:PMC2350904.

37. McLeod D, Mansell J, Harris R, Bailey T, Dowell A, Robson B et al. The collection of patient ethnicity data: A challenge for general practice. *The New Zealand Family Physician* 2000;27(3):51–7.

38. Gill PS, Beavan J, Calvert M and Freemantle N. The unmet need for interpreting provision in UK primary care. *PloS One* 2011;6(6):e20837. doi: 10.1371/journal.pone.0020837. http://dx.doi.org/10.1371/journal.pone.0020837.

39. Sheikh A, Gatrad R and Dhami S. Consultations for people from minority groups. *British Medical Journal (Clinical research ed)*;2008;337(7660):51–3.   http://dx.doi.org/10.1136/bmj.a273. PMCid:PMC2443593.

40. British Medical Association. Ethnicity and first language recording. URL: http://bma.org.uk/practical-support-at-work/contracts/independent-contractors/desethnicity [last accessed 05/08/2013]

41. Bhopal R. Ethnicity, Race, and Health in Multicultural Societies: Foundations for Better Epidemiology, Public Health, and Health Care. Norfolk: Oxford University Press, 2007. http://dx.doi.org/10.1093/acprof:oso/9780198568179.001.0001.