against the use of race or ethnic groups as variables except in research into the impact of racism and disadvantage due to cultural differences on the use and provision of health care.[2] It was specifically to address inequalities that "ethnic" data were included as part of the contract minimum datasets, which provider units in the NHS will have to generate for each patient. The proposed categories are based on those used by the Office of Population Censuses and Surveys—a political and pragmatic classification designed for a census rather than to demarcate sociocultural differences.[14] The chief medical officer is clear that these broad categories may not be sensitive enough to identify important local variations.[15] The availability of such data will, however, lead to a host of papers looking at "ethnicity" and, for example, hospital admissions, length of stay, and use of services. These papers will be limited by the doubtful validity of the variables, their mutability, and their substantial confounding by socioeconomic and educational status—information not recorded in the contract minimum dataset. To discover why different groups have different experiences of health and what can be done to redress the balance we need to disentangle the influences of racism, education, unemployment, and social deprivation.

A thorough investigation of the validity of current classifications is urgently needed; in the meantime, science could be greatly improved by adopting the methodological rigour expected elsewhere in medical research. This would include a discussion of the relevance of collecting data on ethnic group; a description and explanation of the categories, how they are to be used, and the logic underlying them; and the measurement of possible confounders. Any hypothesis should be consistent with the conceptual framework used—whether biological, sociological, or interactional.

In 1991 Bhopal and colleagues called for a wide debate to establish internationally acceptable principles guiding classification and description of ethnic groups.[12] In this issue Senior and Bhopal give nine recommendations on how to make ethnicity a sounder epidemiological variable. If these guidelines were to be followed then new research in this field might start to be comparable and could improve our understanding of differences in health between ethnic groups.

KWAME J McKENZIE
Clinical research fellow

Academic Department of Psychological Medicine,
King's College School of Medicine and Dentistry
and the Institute of Psychiatry,
King's College Hospital,
London SE5 3RS

NATASHA S CROWCROFT
Senior registrar

Department of Public Health Medicine,
Croydon Health Authority,
Croydon CR9 2RH

1 National Library of Medicine. *Medline Silver Platter 3.11 CDROM Medline Express 1989-1.* Bethesda, MD: NLM, 1994.
2 Sheldon TA, Parker H. Race and ethnicity in health research. *J Public Health Med* 1992;**14**: 104-10.
3 Blumenbach JF. *The anthropological treatises of Johann Friedrich Blumenbach.* London: Anthropological Society of London, 1865.
4 Gunarathnam Y. *Health and race. A starting point for managers on improving services for black populations.* London: King's Fund, 1994:ir.
5 Senior PA, Bhopal R. Ethnicity as a variable in epidemiological research. *BMJ* 1994;**309**: 327-30.
6 Ahmad WIU, Sheldon T. "Race" and statistics. In: Ahmad WIU, ed. *The politics of "race" and health.* Bradford: Race Relations Research Unit, University of Bradford, 1992:41-50.
7 Bhopal R. Future research on the health of ethnic minorities; back to basics: a personal view. In: Ahmad WIU, ed. *The politics of "race" and health.* Bradford: Race Relations Research Unit, University of Bradford, 1992:51-2.
8 Fernando S. *Mental health, race and culture* London. Macmillan, 1991.
9 Jones JS. How different are human races? *Nature* 1981;**293**: 188-90.
10 Jenkins R. Social anthropological models of interethnic relations. In: Rex J, Mason D, eds. *Theories of race and ethnic relations.* Cambridge: Cambridge University Press, 1986.
11 Leech K. *A question in dispute: the debate about an "ethnic" question in the Census. Runnymede research report.* London: Runnymede Trust, 1989.
12 Bhopal R, Phillimore P, Kohli HS. Inappropriate use of the term "Asian": an obstacle to ethnicity and health research. *J Public Health Med* 1991;**13**::244-6.
13 Polednak A. Mortality from diabetes mellitus, ischaemic heart disease and cerebrovascular disease among blacks in higher income areas. *Public Health Reports* 1990;**105**:393-9.
14 Sillitoe K. *Asking about race (a summary of the results of field trials of questions about race/ethnicity and other related topics, for the census).* London: Office of Population Censuses and Surveys, 1987.
15 Chief Medical Officer on the State of the Public Health. *Annual report of the chief medical officer of the Department of Health for the year 1992.* London: HMSO, 1993.

Public health and the 1991 census

Non-random underenumeration complicates interpretation

See article on p 327 and editorial on p 286

National censuses provide authoritative data on populations and housing. These results inform public understanding and the governance of society, particularly in respect of strategic planning and the allocation of resources. Assessment of health needs, the purchase and provision of services, equality of access, and evaluation of health outcomes—concerns of providers and purchasers—are also contingent on such information.

The 1991 census broke new ground in two respects, both with major dividends for public health research. It asked respondents whether they had a limiting long term illness or handicap and it asked them to identify their ethnic group.

This last census covered 97.8% of Britain's population—a lower proportion than in previous censuses.[1 2] Comparisons with population estimates based on the 1981 census show that the latest census "missed" some 1.2 million people, half of them young adults but also sizeable numbers of young children and very elderly people. The post-census validation survey located only about a fifth of this shortfall. Census undercoverage in 1991 was not random: it was higher in inner London and other metropolitan areas and among men in their 20s (9% overall and almost 20% in inner London).[1 2] Particular types of households—those in converted or shared accommodation and those in inner London—were also undercounted.[1] These patterns of non-response resemble those of previous censuses and of national social surveys such as the general household survey.[3] Several factors, however, probably contributed to the higher undercount in this census compared with others, such as the increases in one person households and transient populations, census procedures relating to communal establishments, and the unpopularity of the poll tax, leading young people (mainly men) to avoid enumeration.[2-5] Dorling and Simpson illustrate the importance of allowing for undercoverage in particular applications of census data.[4]

To correct for the estimated undercoverage, the Office of Population Censuses and Surveys has developed adjustment factors for inflating the census counts in age-

sex-region categories.[1] However, the unavoidable corollary, that the other characteristics of non-respondents are similar to those of respondents, is clearly precarious. Non-respondents are likely to belong to particular population subgroups—for example, groups with low response rates to government surveys include people without cars, single person households, young socially active people, elderly people living alone, non-professional social classes, unemployed people, those without educational qualifications, and Londoners.[6] If the census missed categories of people traditionally underrepresented in social surveys, then the array of social statistics that the census offers will be correspondingly affected and census based measures of need could understate the true picture. Although correcting for the missing numbers in each age-sex-region category is possible, the data cannot be adjusted to reflect the selective characteristics of those who elected not to participate in the census. Furthermore, an important problem facing census analysts has been the geographical allocation of the underenumeration and the production of population estimates below local authority or health district level. Strategies for allocating the undercount at electoral ward level will be proposed by Diamond and colleagues at next month's meeting of the British Society for Population Studies.

In addition to the undercoverage, the data from the 1991 census are subject to another possible source of bias. Included in the data are almost 900 000 people in households that did not submit census returns because they were "wholly absent"—people whose numbers were estimated by enumerators and whose characteristics have been "imputed" on the basis of other returns.[3] Census records relating to almost 8% of people in inner London were imputed in this way.[3] Although this was probably the safest course of action in the circumstances and an improvement on 1981, when absent households were excluded altogether, concerns over the selectivity of non-respondents still apply.

With the introduction of the question on ethnic group, the 1991 census constituted a historic milestone in Britain's long tradition of decennial censuses, dating back to 1801. The ethnic classification used by the Office of Population Censuses and Surveys was derived on the basis of extensive consultations with ethnic minority organisations and pretesting over a period of two decades.[7,8] The census output is given for 10 ethnic groups, collapsed from an extended classification using 35 categories.[9] Issues relating to the classification of ethnicity and race are discussed elsewhere in this week's journal (p 285, p 327).[10,11]

On the assumption that ethnic minorities are subject to the same level and pattern of undercoverage by age, sex, and region as the rest of the population, the Office of Population Censuses and Surveys has published adjustment factors for inflating their census counts.[1] The weighted factors for non-white groups are higher than those for the white group, as non-white groups are disproportionately represented in younger age groups and inner city areas, categories subject to greater undercoverage. For example, Britain's black males aged 20-29 are estimated to have an undercount of 12-15%, compared with 9% in white males.

In fact, response rates to the 1989 census test and the labour force surveys have been lower among ethnic minorities than in the ethnic majority,[7,12] and the census could have been similarly affected. Preliminary work on sex ratios in the census data, calculated after application of the Office of Population Censuses and Surveys' factors to adjust for undercoverage, shows a shortfall of young male adults in ethnic minority populations, particularly among black people. Neither demography nor migration adequately explains this deficit, and residual undercoverage of males from ethnic minorities therefore seems likely.

Despite this caveat, the 1991 census offers rich opportunities for studying Britain's multiethnic populations. For the first time health authorities have comprehensive data on their ethnic minority populations, which can be used to assess needs and purchase services. Furthermore, ethnic epidemiology, hitherto limited largely to analysis by country of birth and thereby resulting in the omission of about half of the Asian and black populations who are born in Britain, will be revolutionised by the availability of population denominators from the census. No longer can ignorance be used as a cover for inaction or inappropriate action.

The difficulties encountered in the 1991 census will enlighten planning for the census of 2001, which is already being discussed. Censuses provide the most complete and authoritative local profiles of people in Britain and their social conditions—coverage in the 1991 census was in fact similar to that of recent censuses in Australia, Canada, and the United States—and they are subject to greater scrutiny and have higher response rates than any social survey. For these reasons they traditionally serve as the standard for evaluating many other sources of data. The 1991 census provides unique opportunities for the study of the British population, provided that users are aware of the caveats to which the data are subject (publications by the Office of Population Censuses and Surveys and others provide guidance on this) and the likely direction and size of the impact on their results. This challenge should not be missed because of trepidation about census classifications or data quality, because a similar opportunity will not present itself until the twentyfirst century.

We acknowledge comments from Andy Teague, Ian Diamond, and Steve Simpson.

VEENA SONI RALEIGH
Senior research fellow
R BALARAJAN
Director

Institute of Public Health,
University of Surrey,
Guildford GU2 5YL

1 Office of Population Censuses and Surveys. *Undercoverage in Great Britain.* London: HMSO, 1994. (OPCS 1991 census user guide No 58.)
2 Population Statistics Division in collaboration with Census Division OPCS. How complete was the 1991 census? *Population Trends* 1993;71:22-5.
3 Dale A, Marsh C. *The 1991 census user's guide.* London: HMSO, 1993.
4 Dorling D, Simpson S. Those missing millions: implications for social statistics of undercount in the 1991 census. *Radical Statistics* 1993;55:14-35.
5 Diamond I. *Where and who are the "missing million"? Measuring census of population undercount. Regional and local statistics.* Esher: Proceedings of the Statistics Users Council Conference, 1994.
6 Barnes B. Studies of non-respondents in OPCS household surveys, using census data: past studies and future plans. OPCS Social Survey Division. *Survey Methodology Bulletin* 1992;30:21-5.
7 White P. A question on ethnic group for the census: findings from the 1989 census test post-enumeration survey. *Population Trends* 1990;59:11-20.
8 Sillitoe K, White P. Ethnic group and the British census: the search for a question. *Journal of the Royal Statistical Society* 1992;155:141-63.
9 Teague A. Ethnic group: first results from the 1991 census. *Population Trends* 1993;72:12-7.
10 McKenzie K, Crowcroft N. Race, ethnicity, culture, and science. *BMJ* 1994;309:285-6.
11 Senior PA, Bhopal R. Ethnicity as a variable in epidemiological research. *BMJ* 1994;309: 327-30.
12 Owen C. Using the labour force survey to estimate Britain's ethnic minority populations. *Population Trends* 1993;72:18-23.