

# CeLSIUS News

Issue 7

May 2006



## Come and see us!

We're going to the meetings below - if you're going too, we'd be pleased to meet you:

19 – 21 June

Third International Conference on Population Geographies,  
University of Liverpool

There will be a CeLSIUS exhibition stand and Chris Marshall will be giving a presentation on 20 June at 3.50 pm

21 – 23 June

European Population Conference  
University of Liverpool

Emily Grundy will be giving a presentation based on analysis of the LS

12 - 14 July

MOLS 2006,  
University of Essex

Julian Buxton will be giving a presentation on 12 July at 11 am

17-20 July

ESRC Research Methods Festival,  
University of Oxford

There will be a Census of Population Programme exhibition stand and Emily Grundy will be giving a presentation during the Resources session on 18 July at 9.15 am

12 - 14 September

International Conference On Child Cohort Studies,  
University of Oxford

Lynda Clarke will be giving a presentation (T.B.A.)

## Great news – CeLSIUS to continue until 2011!

Welcome to Issue 7 of CeLSIUS News. We are pleased to announce that we have accepted the offer of a research grant from ESRC that will enable us to continue our work until the end of July 2011. We shall continue to provide expert research and computing support to LS users as well as produce more online training modules to cover particular areas of the LS, hold workshops on different aspects of the data and provide 'off the peg' data for downloading straight from the web, which will speed up access to standardised data. We shall also be contributing to an integrated web portal and dissemination service that will bring together all the Census Programme data units. Look out for news of our activities in future editions of CeLSIUS News.

We have just finished 'version one' of an online training module on geography data in the LS – this has now been sent off to ONS for clearance and will be launched at the Population Geographies conference in Liverpool in June (see left for further details). It will be available from our web site then too. Work has begun on a module on regression based approaches to LS data analysis – watch this space for further news.

We are always keen to spread the word about the ONS LS and the CeLSIUS service, so please help us to do this by passing on the link to CeLSIUS news to your colleagues in the UK academic community.

### In this issue:

LS research opportunities  
Claimant count data soon to be available  
Revised procedures for accessing the LS  
New online training module  
News from ONS  
"The joys and challenges of attaching area data to the ONS Longitudinal Study" by Paul Norman  
Latest LS publications

If you wish to comment on this issue or contribute an article to future editions, please contact: [celsius@lshtm.ac.uk](mailto:celsius@lshtm.ac.uk).

CeLSIUS, Centre for Population Studies, London School of Hygiene & Tropical Medicine, 49-51 Bedford Square, London WC1B 3DP. Tel: 020 7299 4634. Email: [celsius@lshtm.ac.uk](mailto:celsius@lshtm.ac.uk).  
Web: [www.celsius.lshtm.ac.uk](http://www.celsius.lshtm.ac.uk).

To join our free e-discussion list, go to [www.jiscmail.ac.uk/lists/ls-support.html](http://www.jiscmail.ac.uk/lists/ls-support.html).

## LS research opportunities

The LS is a data set comprising linked census and event records for 1% of the population of England and Wales (about 500,000 people at any one census). Use of the LS and the support we provide to staff and students of UK academic institutions is **absolutely free**. 2001 Census data was added to the LS in September 2004, providing 30 years' follow-up of original sample members. This has opened up extensive new research opportunities, especially as the 2001 Census included new or improved questions on education, self-rated health, relationships within the household, care-giving, and religion. The CeLSIUS web site has details of all the current and recent research done using the LS including lists of publications but the data has great potential in many other areas as well. For example:

- o Examining links between childhood and adult circumstances, including health, among those age 30-46 in 2001 (who would have been aged 15 or under in 1971).
- o Tracking change through mid and later life for those aged 40 or more in 1971.
- o Intergenerational comparison (e.g. analyses of family building patterns by family type in childhood, and by fertility history of mother for those aged under 15 in 1971).

For further information about using the LS please email [celsius@lshtm.ac.uk](mailto:celsius@lshtm.ac.uk), visit [www.celsius.lshtm.ac.uk](http://www.celsius.lshtm.ac.uk), or call Jo Tomlinson on 020 7299 4634.

## Claimant count data soon to be available

ONS has linked unemployment claimant count data to the Longitudinal Study. The new information provides a continuous record of LS members' unemployment events from 1994 to 2004. Linkage to the LS permits analysis of unemployment spells in the context of, for example, 1991 and 2001 characteristics and life events during the 1990s.

Initial testing of the data is nearing completion and ONS are currently constructing 'CCC quality tables' with the number and length of unemployment spells by various socio-demographic characteristics.

Details of the linkage (including the matching of claimants to LS members of over 97 per cent) will be described at the Methodology of Longitudinal Surveys conference at Essex University, 12-14 July. (CeLSIUS will also be giving a presentation there, for more details see [www.iser.essex.ac.uk/ulsc/mols2006](http://www.iser.essex.ac.uk/ulsc/mols2006)). This work will form part of CCC metadata which will be made available from the ONS web site along with a User Guide, to accompany release of the data.

As happened with the LS-2001 Census data link, an invitation to submit beta test applications was circulated to academics via the Celsius web site, along with draft metadata produced by ONS. Applications were submitted to the 25 May meeting of the LS Research Board and testing is due to begin in July 2006. The aim is to comprehensively test the accuracy, coherence and comparability of the data, as well as any potential disclosivity implications before its eventual release later this year.

To receive an email alert when the claimant data becomes available, please join our free e-discussion list at [www.jiscmail.ac.uk/lists/ls-support.html](http://www.jiscmail.ac.uk/lists/ls-support.html), or look out for an announcement on our web site.

## Revised Procedures for Accessing the LS

ONS recently reviewed how users access their microdata and the review panel made some recommendations. Following on from these the ONS LS Development Team have revised how users may access the LS. For full details please visit [www.celsius.lshtm.ac.uk/news.html](http://www.celsius.lshtm.ac.uk/news.html).

## New online training module

A new geography module will be available shortly from the LS training web page: [www.celsius.lshtm.ac.uk/training.html](http://www.celsius.lshtm.ac.uk/training.html). Chris will be featuring this module in his presentation at the Third International Conference on Population Geographies at the University of Liverpool on 20 June. For further details of this conference, go to [www.geog.leeds.ac.uk/groups/pgrg/conf2006.html](http://www.geog.leeds.ac.uk/groups/pgrg/conf2006.html).

## News from ONS

### New chair of LS Research Board:

ONS welcomes Joy Dobbs as the new chair of LSRB, and would like to thank Peter Goldblatt for his many years of service as Chair.

### 2006 Data Refresh:

The latest events data refresh is on schedule for completion by the end of July 2006. This will take LS births and deaths information to the end of 2004, cancers to the end of 2003 and other events to Spring 2006. The refreshed tables have already been successfully transferred from M204 to SQL in Titchfield and most are currently in London for validation and comparison testing. Also, three new social class variables have been added to the 1991 and 2001 Census tables:

Registrar General's Social Class 2001 using David Rose's approximation based on NS-SEC (ROSERG0).

Registrar General's social class in 2001 using LS SOC90 codes and the closed derivation matrix supplied by the ONS Occupational Information Unit (SCLAS0).

National Statistics Socio-economic classification 1991 based on SOC90 coding, coded for consistency with 2001 NS-SEC LS data (NS-SEC90).

These fields have been alpha tested and accompanying metadata will be included in the data dictionary.

### LS Open Access Area:

The LS Development team has recently moved to new accommodation on the 2nd floor of the main ONS Drummond Gate building. This offers improved space and facilities for the Open Access Area. It has also been possible to dispense with the locks which restricted access to the Bessborough safe setting.

This move means that CeLSIUS Support Officers, Julian and Chris, are also now located on the second floor when at ONS. Their ONS telephone number remains the same: 020 7533 5638.

## The joys and challenges of attaching area data to the ONS Longitudinal Study

Paul Norman CCSR, University of Manchester, [paul.norman@manchester.ac.uk](mailto:paul.norman@manchester.ac.uk)

I am currently involved in my second project using the LS in which I have used area data attached to LS members' records. The 'attaching data' stage of the first project occurred whilst LS academics were supported by the Centre for Longitudinal Study with the donkey work carried out by Kevin Lynch and Sarah Jones. For my current project, support is provided by Chris Marshall of CeLSIUS. I am grateful to them all for their help and patience.

Analysing longitudinal socio-demographic microdata linked to area data has massive research potential, but presents various inter-related conceptual and practical problems. My intention here is to reduce the possibility of people running into problems before they occur rather than to put people off! To give this article some context, first I am going to describe two longitudinal analysis frameworks involving both individual and area data.

- i) The outcome for an individual can be modelled in relation to that person's characteristics and contextual data about specific areas and/or types of areas in which that person lived at various time points.
- ii) Individual level data are aggregated to geographical areas or types of area at different time points. If changes are found to occur for areas, can these be explained by individuals changing their characteristics and/or their locations?

There are other approaches, please forgive me if I haven't described your favourite!

The ONS LS has many time points at which information about each member's location is known. In simple terms, these relate to 1939 (national registration), census years since 1971 and the location of demographic events. Theoretically, you can attach information about the areas in which people were living whenever their location is known. In practice, if you go down this route, you will need to consider various aspects.

*What geography? Specific locations ...* Ideally, you select your favourite geographical scale, obtain some data about these areas, send the files to your CeLSIUS support officer who will attach your data to LS members and soon you will be advancing the research frontiers. Whilst this may emulate good science and provide application-relevant results, life is not so simple! Your first stop is the LS Data Dictionary. Here, perusal of keywords reveals many variables under 'Geog' and prefixed 'Small-area'. Numerous different geographies are available and for different years. These geographies mainly relate LS members to health and census or administrative geographies. Some are available for analysis, but many are not. If you are lucky, you might find a geography appropriate to your proposed research available at all time points, but there's a good chance you won't. Recent geographies are more likely to suit your application, but if you have an interest in the influence of time, previous geographies are important. You will almost certainly have to compromise.

An important consideration is geographical consistency over time. Without reasonable consistency you cannot objectively assess inter-relationships and relative influence of individual and area change. Consider aggregating electoral wards into a custom geography. Bear in mind, however, that ward geography changed dramatically between 1971 and 2001 and that the derivation of a time-series of lookup tables of wards to current 'Primary Care Trusts', for example, will challenge even the most adept Su Doku solvers! Alternatively for a regional analysis, post-1974 counties can be nested into the current Government Office Regions. Note that microdata with named locations attached may not be released by ONS for confidentiality reasons, but this may not preclude your analysis from taking place with results available for your geography.

*... or area types?* Rather than specific (i.e. named and identifiable) locations, another approach is to carry out analyses using generic descriptions or categorisations attached to LS members' records. Thus your analysis could include, for the location of residence per time point, the deprivation quintile, an urban/rural indicator or a geodemographic description. This slant produces results which lend themselves to digestible descriptive reporting, for example, "people living for three consecutive censuses in relatively non-deprived areas have lower risk of limiting long-term illness than people living in deprived areas for the same time period". Moreover, since a classification may have a limited number of categorisations, you decrease the risk of breaching confidentiality restrictions. Remember, however, that any scheme is time and geography specific and may not be consistently available for all LS time points. Degree of urban-ness or rurality may inform the relationship between age, migration patterns and (counter-) urbanisation, but as variables may have been differently defined after each census, you may not be comparing like with like. For my latest LS project I have calculated 1971, 1981, 1991 and 2001 ward population densities based on persons per hectare. I acknowledge that this may be less than ideal as an urban/rural indicator, but the results will be underpinned by consistency.

Attaching area data to the LS has huge potential to add value to analysis. However, many of the LS geographical variables are dated and possibly redundant and there are linkage difficulties because

different area reference codes have been used for the same geography by different organisations in the past. This is not to blame any organisation, just a result of degradation through time and only relatively recent awareness of the situation and an increasing desire for time-series analyses. Various academics and ONS geography have been encouraging data disseminators to adopt the same area naming and alphanumeric coding systems for census and administrative geographies. Whilst much progress has been made recently, we are still stuck with inconsistencies for previous geographies. Before you devise area-based measures which you want to use in your LS project, get in touch with CeLSIUS and ask for a lookup table with names and codes of the geography you want to use, even for census data.

*So what can we do?* Frankly I'm not sure, but for a start I would like to explore the possibility of trying to attach a grid reference to LS members based on the most detailed level of geographic locator available on each record. This would enable linkage with the geography of research interest but inevitably information will vary in quality with older records likely to be least reliable. However, linking LS records with measures derived from the 'Linking Censuses Through Time' project or with deprivation scores I have calculated for 1971 through to 2001 all based on the 2001 ward definitions increases research potential.

Note that the views above have been expressed by a grumpy old geographer who is looking forward to being corrected by Julian Buxton and Chris Marshall's paper on 'Geography and Geographical Analysis using the ONS LS' at the International Population Geography Conference in Liverpool this June.

## Latest LS publications

*Fertility history and health in later life: a record linkage study in England and Wales.* E. Grundy and C. Tomassini (2005). *Social Science and Medicine* 61: 217-228

*Migration and mobility: the life chances of Britain's minority ethnic communities.* L. Platt (2005). A report from the Joseph Rowntree Foundation-funded project: The Intergenerational Social Mobility of Minority Ethnic Groups, February to December 2004. Available from the Joseph Rowntree Foundation web site: [www.jrf.org.uk](http://www.jrf.org.uk).

*New destinations? Assessing the post-migration social mobility of minority ethnic groups in England and Wales.* L. Platt (2005). *Social and Policy Administration* 39 (6): 697-721