

# ESRC Longitudinal Studies Review 2017

## Further analysis of responses to the consultation

### Paper 3:

### The capacity and infrastructure for longitudinal research

April 2017

### Report author:

Ruth Townsley, Independent Researcher

The views represented in this report are from those who responded to the consultation and do not represent the views of ESRC



## Introduction

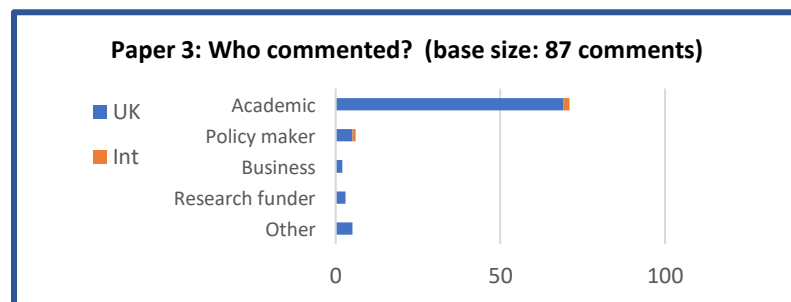
The ESRC Longitudinal Studies Review 2017 is exploring the current and future scientific and policy-relevant need for longitudinal research resources. The review started in October 2016 and will report to ESRC Council early in 2018.

An open online consultation in autumn 2016 sought input broadly, resulting in 637 completed responses from UK (83.4%) and international (16.6%) respondents. Respondents were predominantly from the academic sector (80%) as well as government, civil society and business sectors (20%).

The main findings of the consultation were published in December 2016 in an [initial report](#). This report is supplemented by short briefing papers that examine key themes from the consultation data in more detail.

## Briefing paper 3: The capacity and infrastructure for longitudinal research

This paper examines respondents' comments about the capacity and infrastructure for longitudinal research in the UK. A wildcard search using the strings train\*, skill\* and capacity was conducted, alongside further analysis of existing coded comments relating to infrastructure and capacity building, which identified a total of 87 relevant comments and suggestions.



## Identification of skills gaps

Respondents raised concerns about the continued need for specialist expertise in key areas of longitudinal research and highlighted potential skills gaps and shortages of trained individuals in the following areas:

- > Designing, leading and managing longitudinal studies
- > Analysis of longitudinal data – data analysts who can analyse longitudinal data appropriately and effectively in order to extract maximum value and use from existing datasets; this includes knowledge and understanding of approaches to handling missing data, weighting and multiple imputation, addressing sources of bias to aid causal inference, and advanced statistical techniques
- > Awareness, understanding and analytic skills needed to consider and fully exploit possibilities for cross-study analysis
- > Awareness, understanding and analytic skills for the integration of biosocial resources
- > Use of longitudinal data – frequent and infrequent data users from academic, policy, business and third sector organisations who have sufficient training and expertise to analyse and model data effectively: not doing so may lead to erroneous inferences for policy and science.

Several respondents suggested that the UK is 'lagging behind' other European countries and the US in terms of wider capacity, across all sectors, to analyse and make use of longitudinal sociological data.

"The skills required to lead a study, design a study and analyse the data are not discussed and remain under-valued and in deficit in the U.K. Linking skills and methods and data innovation and investment is an important consideration for future longitudinal studies." (ID 79)

## Actions needed to build capacity

Respondents made many suggestions about how best to build capacity to address the skills gaps outlined above. These included:

- > Training to ensure there is a sufficient on-going capacity within the next generation - including training on designing, managing and leading longitudinal studies; and on analytic techniques for handling missing data
- > Wider access to training for non-academic users (including policy makers) to improve the effective use and potential impact of longitudinal resources
- > Considering developing common tools, appropriate software and good practice guidance on statistical methods for managing missing data
- > Maintaining the range of training opportunities introduced since the 2006 LS review (Q-Step, AQM enhanced stipend, Soc-B Centre for Doctoral Training in Biosocial Research) and supporting new training initiatives (e.g. the range of training resources on analytic techniques being planned by the Centre for Longitudinal Studies) to continue to improve skills in key areas
- > Continued support (through CLOSER and other initiatives) for cross-study and cross-cohort analysis, both in maximising and developing existing resources, and in building capacity to undertake new analysis, e.g. through training, resources, and awareness-raising to fully exploit the potential of existing longitudinal data
- > Continued support for capacity development (designing studies, collecting data, analysis) in biosocial research, e.g. through the Soc-B Centre, but also more widely through other training initiatives, including at Masters level, for other data analysts and users, and potentially to enable interviewers (rather than nurses) to collect biometric measures so creating scope for collecting new data within existing studies.

"There is still a need to enhance the capacity of social scientists to use longitudinal data. There are hardly any issues that cannot be addressed without the use of such data - and many that can't without." (ID 213)

"A challenge for longitudinal studies is the lack of hands on training on the appropriate use of recent methodological advances in missing data handling in realistic analytic scenarios. Courses with accessible material as well as wider dissemination of how principled approaches that deal with missing data bias are appropriately applied in complex multivariate longitudinal analysis is needed." (ID 81)

"The suite of longitudinal surveys in the UK is world-leading. The ESRC is the only funder able to support this long-term investment in the social science data infrastructure for the UK research community. Much still needs to be done to train researchers in using these data sources and a focus on teaching quantitative analysis techniques through initiatives such as Q-Step should be maintained." (ID 430)

## Supporting the UK infrastructure for longitudinal research

Many respondents stated the need for a permanent infrastructure to support and protect the UK's suite of large scale, population based longitudinal studies. Intrinsic to this is an understanding of the unique social and scientific value of current longitudinal resources and the importance of maintaining them as a national asset for future generations. Respondents highlighted a number of areas where infrastructure support is needed, or could be further developed:

- > Maintaining and curating existing datasets and ensuring these remain useable in the future

- > Supporting continued population representation based on probability sampling and ensuring that studies remain representative of their target population – investing in participant engagement through evidence of what works; other methods and best practice to maximise participant response; capacity building and application of statistical methods to deal with missing data
- > Continued and increased resource for statistical and methodological research and innovation alongside and within existing studies – these include developments in data record linkage, cross-cohort analysis and integration of biomedical data – links to health, tax and educational records and comparison across UK birth cohorts are informative but methodologically challenging and require dedicated resource to achieve effective outcomes
- > Continued support for biosocial infrastructure investments (including nurse panel development) to enable integration of biomedical data within national longitudinal studies
- > Enabling new research from these infrastructural achievements through grant funding programmes – includes funding new research from biosocial data developments; and interdisciplinary research which makes innovative use of new sources of data to answer (new) policy questions
- > More recognition of the value of longitudinal research skills and capacity – for example, acknowledging that trained and expert field-based staff are a national resource which should be supported and maintained; and ensuring that those who design, run and analyse longitudinal studies/data are more appropriately recognised within HEI reward and promotion systems.
- > The importance of sustaining reliable censuses without which there are no baselines for sample surveys and other studies
- > Reconsidering the case for a new birth cohort – this was felt by some to be of critical importance for sustaining the UK's longitudinal research capacity – maintaining a continuous flow of data is needed to track societal changes and their impacts on generations born in the 21<sup>st</sup> century – the current lack of sequence since the MCS presents a gap in scientific knowledge and limits the extent to which some policy questions can be addressed.

"It's of vital importance that longitudinal surveys remain representative of their target population as they provide reliable insights on topics not covered by administrative and other forms of 'Big Data', but also act like a gold standard where the generalisability of findings from 'Big Data' and other unbalanced samples can be empirically verified." (ID 81)

"UK Data Archive and ESRC's policies to put funded data into the public domain are excellent. Continued funding is needed so that existing data remains usable in the future and new longitudinal data can be collected." (ID 329)

"I would welcome anything that enables us to give our interviewers and nurses confidence that longitudinal survey research is valued and will be an essential part of social research in the long term. Many interviewers who have been with us for many years are unclear of their future. A strategy to work with the agencies to capacity-build and set out how the interviewer and nurse roles can most effectively evolve would be very welcome." (ID 15)

"The new birth cohort study (Life Study) recommended by the 2006 review has closed - but the need for longitudinal data on generations born in the 21<sup>st</sup> century becomes ever more pressing. Child and adult life courses are changing rapidly - and are becoming more unequal - yet we lack the sequence of cohort studies (1946, 1958, 1970, 2000/01) on which science and evidence-informed policy has previously relied." (ID 270)