

Millennium Cohort Study Fourth Survey

A Guide to the Linked Education Administrative Datasets

First Edition

Jon Johnson and Rachel Rosenberg

July 2011

Centre for Longitudinal Studies
Institute of Education, University of London



First published July 2011
by the Centre for Longitudinal Studies
Institute of Education, University of London
20 Bedford Way
London WC1H 0AL

Website: www.cls.ioe.ac.uk

© Centre for Longitudinal Studies

The Centre for Longitudinal Studies (CLS) is an Economic and Social Research Council Resource Centre and is devoted to the collection, management and analysis of large-scale longitudinal data. It is based at the Institute of Education, University of London. The Centre houses three internationally-renowned birth cohort studies: the 1958 National Child Development Study (NCDS), the 1970 British Cohort Study (BCS) and the Millennium Cohort Study (MCS).

The views expressed in this work are those of the authors and do not necessarily reflect the views of the Economic and Social Research Council. All errors and omissions remain those of the authors.

Contents

ACKNOWLEDGEMENTS 4

PREFACE..... 4

BACKGROUND..... 5

1. Introduction..... 5

2. Consent to Education Data Linkage 6

 Table 1: Consent Outcomes 6

3. Education Data Linkage..... 6

3. Linkage outcome..... 7

 Table 2: Linkage Outcome..... 7

3. Data variables..... 7

 Table 3: Data Variables (England) 8

ACKNOWLEDGEMENTS

We are grateful to the Department for Education for providing linkage services.

PREFACE

This document has been prepared to accompany the deposit of data from the fourth survey of the Millennium Cohort Study at age 7 (MCS4) with the UK Data Archive at the University of Essex.

The first edition of the Guide to the Linked Education data covers linkage for schools in England. Linkage for Scotland is forthcoming and that for Wales and Northern Ireland is pending negotiations with the respective Assemblies.

BACKGROUND

1. Introduction

The Millennium Cohort Study (MCS) is the fourth of Britain's world-renowned national longitudinal birth cohort studies. Each follows a large sample of individuals born over a limited period of time through the course of their lives, charting the effects of events and circumstances in early life on outcomes and achievements later on. They show how histories of health, wealth, education, family and employment are interwoven for individuals and vary between them.

The MCS Study offers large-scale information about the New Century's children, and the families who are bringing them up, for the four countries of the United Kingdom. Its first sweep, carried out during 2001-2002, laid the foundations for a major new longitudinal research resource, taking a new "year long" cohort of around 19,000 babies. In England and Wales they were born over the 12 months starting in September 2000, and over 13½ months from late November 2000 in Scotland and Northern Ireland. Information was collected from parents when the babies were aged nine months. The sample design allowed for disproportionate representation of families living in areas of child poverty, and in areas of England with high ethnic minority populations. The first survey recorded the circumstances of pregnancy and birth, as well as those of the all-important early months of life, and the social and economic background of the family into which the children have been born.

This multidisciplinary baseline data reveal the diversity of starting points from which these 'Children of the New Century' have set out. The second survey data add a longitudinal element to the dataset and will allow researchers for the first time to chart the changing circumstances of children and their families, and relate age 3 outcomes to earlier circumstances and experiences. The third survey visited the families when the children reached age 5 and had started primary school.

The fourth returned to the families in 2008 when the children were age 7 and collected information from parents and carer's about schooling, the child's development and parental assessments of cognitive ability and behaviour. This data also included self reported information about the school.

Further details of the data available from the main survey are available from the CLS website and in particular the *MCS Guide to the Datasets* at www.cls.ioe.ac.uk/mcs.

2. Consent to Education Data Linkage

During the fourth sweep, the parents / carer's were asked for consent to link to the child's education records. Details of the form and the wording of the consent is reproduced in the MCS4 Technical Report Appendices Section A16, consent form 3.

Table 1: Consent Outcomes

Country of interview	Yes	No	No form	Total
England	8447	293	259	8999
Wales	1898	67	30	1995
Scotland	1536	62	51	1649
Northern Ireland	1288	73	38	1399
Total	13169	495	378	14042

3. Education Data Linkage

During the fourth sweep, the parents / carer's were asked the name of the school the cohort child was attending was recorded and a lookup table was used by the field work agency to confirm the name and location of the school. This produced a dataset which included the URN of the school, the Local Education Authority and the school number used by the Department of Education in England, the Welsh Assembly, the Scottish Parliament and NISRA in Northern Ireland.

England

After some cleaning at CLS, to correct incorrectly assigned schools (normally the name of a previous school which has since closed and renamed). CLS prepared a file 8444 cases to the Department for Education in England, (this excluded 3 cases where the interview was in England but the school reported was in another country.

The file contained the following information on the cohort child:

Name (first and last)
Sex
Date of Birth
Postcode of residence at interview
Name of school
URN of school of current school
Local Education Authority of current school
School Number of current school

The NPD database team matched the data on the following basis.

- a. Direct match on all variables
- b. Direct match of 4 of the 5 variables or 3 of 4 etc

- c. Fuzzy matches on names – usually that the first 3 or 4 characters are the same – again all manually checked.
- d. Any other combinations we think are safe based on the data supplied

There is unfortunately not a variable that indicates the method used.

3. Linkage outcome

Table 2 shows the outcome of the linkages carried out.

Table 2: Linkage Outcome

	England	Wales	Scotland	N. Ireland
No of consents	8447	1898	1536	1288
No of cases sent for linkage	8444	n/a	n/a	n/a
No. of cases successfully linked	6841	n/a	n/a	n/a
% of cases successfully linked	81%	n/a	n/a	n/a

3. Data variables

Table 3 lists the variables obtained from the linkage to the National Pupil Database (NPD).

Those labelled KS1_ are from the KS1 data held in the NPD. The remaining variables are from the school census 2010. In addition to these are LEANON and SCHOOLANON, which are anonymised Local Education Authority and School Number, to allow comparison of results across LEA and school.

Table 3: Data Variables (England)

File name: mcs4_ks1_plasc.sav

MCSID	MCS Research ID
CNUM	Cohort Child Number
ACADYR	KS1_ACADYR Academic year
AGESTART	KS1_AGE_START Age of pupil at start of the academic year (in full years)
TOECODE	KS1_ToE_CODE Type of establishment code taken from Edubase
NFTYPE	KS1_NFTYPE Institution type generated by AAT - numeric
MMSCH	KS1_MMSCH Maintained Mainstream school (including CTCs and Academies)
MMSCH2	KS1_MMSCH2 Maintained Mainstream school (excluding CTCs and Academies)
MSCH	KS1_MSCH Maintained school (including CTCs and Academies)
MSCH2	KS1_MSCH2 Maintained school (excluding CTCs and Academies)
MOB1	KS1_MOB1 Pupil joined in latest academic year (year 2)
MOB2	KS1_MOB2 Pupil joined in previous academic year (year 1)
APS	KS1_APS Average attainment point score
SPKLISTN	KS1_SPEAKANDLISTEN National Curriculum level (Teacher Assessment)
LEVLSPLI	KS1_LEVLSPLI Achieved Level 1 or below in KS1 Speaking + Listening
EVXSPLI	KS1_LEVXSPLI Achieved Level 2 or above in KS1 Speaking + Listening
EVAXSPLI	KS1_LEVAXSPLI Achieved Level 3 or above in KS1 Speaking + Listening
READING	KS1_READING National Curriculum level awarded for reading
READPNTS	KS1_READPOINTS Reading attainment point score
EVLREAD	KS1_LEVLREAD Achieved Level 1 or below in KS1 Reading
LEVXREAD	KS1_LEVXREAD Achieved Level 2 or above in KS1 Reading
LV2BREAD	KS1_LEV2BREAD Achieved Level 2B or above in KS1 Reading
LVAXREAD	KS1_LEVAXREA Achieved Level 3 or above in KS1 Reading
WRITING	KS1_WRITING National Curriculum level awarded for writing
WRITPNTS	KS1_WRITPOINTS Writing attainment point score
LVLWRIT	KS1_LEVLWRIT Achieved Level 1 or below in KS1 Writing
LVXWRIT	KS1_LEVXWRIT Achieved Level 2 or above in KS1 Writing
LV2BWRIT	KS1_LEV2BWRIT Achieved Level 2B or above in KS1 Writing
LVAXWRIT	KS1_LEVAXWRIT Achieved Level 3 or above in KS1 Writing
RDWRIT	KS1_READWRIT English Average Level (derived from Reading and Writing)
RDWRPNTS	KS1_READWRITPOINTS Reading and writing attainment point score
MATHS	KS1_MATHS National Curriculum level awarded for Maths
MPOINTS	KS1_MATPOINTS Maths attainment point score
LVLMAT	KS1_LEVLMAT Achieved Level 1 or below in KS1 Maths
LVXMAT	KS1_LEVXMAT Achieved Level 2 or above in KS1 Maths
LV2BMAT	KS1_LEV2BMAT Achieved Level 2B or above in KS1 Maths
LVAXMAT	KS1_LEVAXMAT Achieved Level 3 or above in KS1 Maths
SCXPINVT	KS1_SCIEXPINVEST National Curriculum level awarded for scientific enquiry
SCLIFPRT	KS1_SCILIFPROT National Curriculum level awarded for life and living processes
SCMPROP	KS1_SCIMATPROP National Curriculum level awarded for materials and properties
SCIPPROC	KS1_SCIPHYSPROC National Curriculum level awarded for physical processes

SCIENCE	KS1_SCIENCE National Curriculum level awarded for overall Science
SCPNTS	KS1_SCIPOINTS Overall Science attainment point score
LVLSCI	KS1_LEVLSCI Achieved Level 1 or below in KS1 overall Science
LVXSCI	KS1_LEVXSCI Achieved Level 2 or above in KS1 overall Science
LVAXSCI	KS1_LEVAXSCI Achieved Level 3 or above in KS1 overall Science
LEAANON	LEA / Board anonymised
SCHOOLANON	School no, anonymised
RECSTAT	Record Status
ACAYEAR	AcademicYear
CENTERM	CensusTerm
ONROLL	OnRoll
ETHNMIN	Pupils minor ethnic group based on ethnic code
ETHNMAJ	Pupils major ethnic group based on ethnic code
ETHNSRC	Source of pupils ethnicity code
FSMELIG	Pupil eligible for free school meals
CONNEX	Do parents consent to pupil data being shared with Connexions
LANGMIN	Pupils minor language group based on language code
LANGMAJ	Pupils major language group based on language code
GANDTIND	Indicates if the pupil is in the Gifted and Talented cohort for the school
MTRAVEL	Indicates the usual mode of travel used by the pupil
ENRLSTAT	Indicates enrolment status of pupil
TYPCLASS	Type of Class
ENTRYDTE	Date of entry to current school
PARTTIME	Indicates if the pupil is part time or not
BOARDER	Indicates if the pupil is a boarder or not
NCYRACT	The year group in which the pupil is taught
SENPROV	Provision types under the SEN Code of Practice
SENPROV2	Pupils major SEN provision group based on SEN provision code
SENUIND	Pupil with SEN in a mainstream school is a member of a SEN unit
RPROVIND	Pupil with SEN in a mainstream school is a member of a resourced provision
MOBIND	Pupils postcode has changed from previous Spring Census (Spring only)
DCURRSCH	Distance in miles from pupils postcode to current school (Spring only)
DNEARSCH	Distance in miles from pupils postcode to nearest suitable school (Spring)