



Opportunities for Statistical Data Linkage in the UK

Paul Allin

Ben Humberstone

Introduction: Opportunities for Statistical Data Linkage in the UK

- Increasing awareness of the importance of administrative data and the benefits of using it including increased efficiency across Government and the potential for significant savings
- Data linkage is the key to unlocking these benefits.
- Examine developments in the UK, specifically the Citizen's Information Project and the Integrated Population Statistics System.

Citizen Information Project: the vision

- A shared national infrastructure of basic personal data
- to facilitate major economies, efficiencies and service improvements in the UK public sector.

Started: Summer 2002

Ends: June 05 end-stage report to Ministers with recommendation on implementation

The size of the challenge

- 59.2 million people (ONS, 2002)
- 7-10 million changes of individual contact details a year
- increasing mobility including international migration (in & out)
- diversity in living arrangements
- thousands of registers & databases
- major developments already under way

CIP: mid-stage assessment

- potential efficiency benefits confirmed
- range of potential options
- better through NIR than separate register
- needs further examination
- opportunities for adding value to existing database developments in shorter term
- need to meet efficiency and analytical requirements that cover whole population, rather than just adults

Current assumptions

- will be National Identity Register with
 - name and alias
 - one or more residential addresses
 - date and place of birth
 - sex
 - date of death
 - unique reference number (NIR No.)
 - social security number (NINO)

Current assumptions

- coverage of ID Cards scheme:
 - people aged 16 or over
 - legally resident in UK for 3 months or more
- NIR will be:
 - built from empty
 - virtually complete by around 2020, assuming ID cards will become compulsory
 - notified of deaths in UK by register offices
- data sharing provisions

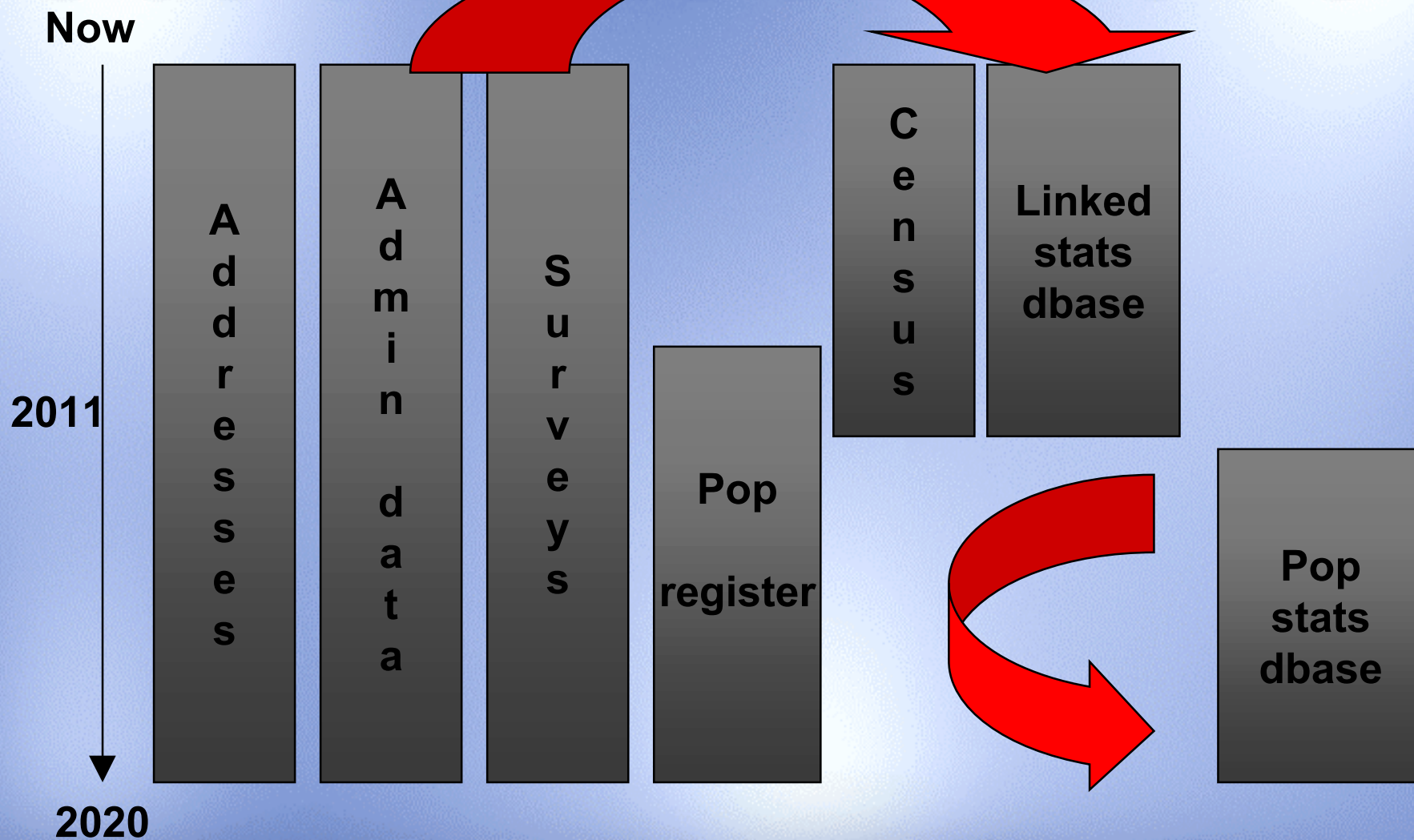
Population coverage issues

- Adults & children
- Usually resident e.g. EEA nationals working in UK
- Main residence?
- Different population bases
- Individuals, households & families
- UK nationals living overseas
- Emigration from UK (means?)

CIP, statistics & research

- 2011 and future censuses
- integrated population statistics system
- various population bases
- migration - defining and measuring
- secure data sharing for statistics & research

Towards an integrated population statistics system



IPSS drivers

- Increased demand for more frequent information at small-area level
- Falling response rates to surveys and censuses
- Government reviews and policy:
 - Allsopp Report: Review of Statistics for Economic Policymaking
 - Neighbourhood Renewal, Neighbourhood Statistics (NeSS)

Data sharing: opportunity or threat?

- Administrative data: part of the national data strategy?
- Establishing data requirements
- Learning from the pros and cons of secondary analysis of survey data
- Opportunities for longitudinal research
- Maintaining a balance of data resources

Key challenges to construction of IPSS

- Legality of linkage - creation of a secure data matching environment
- Political and public acceptability
- Quality of admin data
- Dependencies on other projects (population/address register)
- Confidentiality and disclosure
- Finance

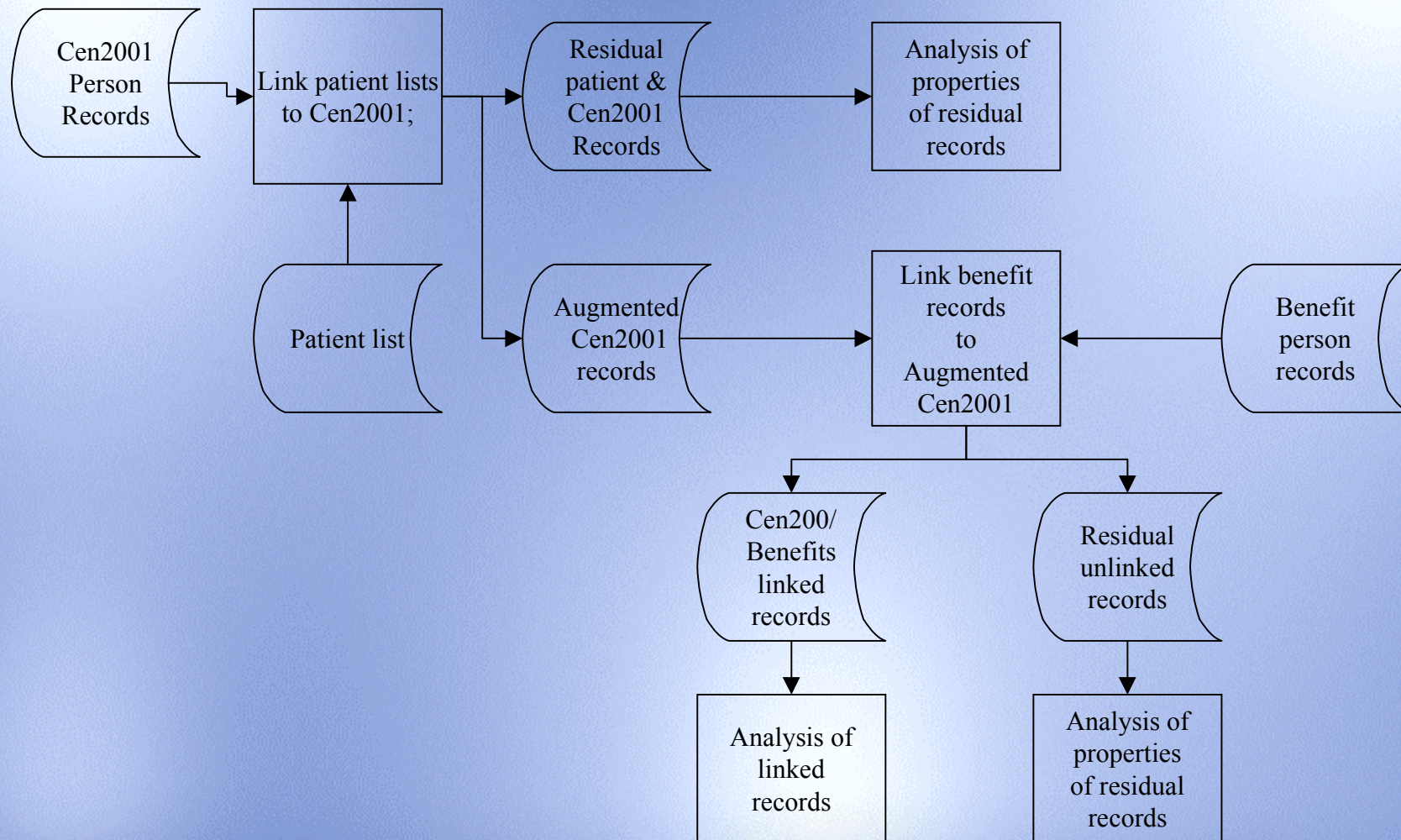
IPSS feasibility study

- Key objectives:
 - Test the legal and policy issues
 - Develop and evaluate a test IPSS by 2007
 - Establish practicality of linking administrative data drawn from different sources based on the premise of an IPSS
 - Establish the statistical benefits of an IPSS database to ONS and across Government
 - Contribute to “Go/No Go” decision by 2007

IPSS feasibility study

- Match a sample of data from three reliable sources:
 - Benefit data
 - Patient lists
 - 2001 Census

IPSS: High level feasibility project flowchart



Evaluation

- Relative quality of the different datasets
- Are link rates sufficiently high between datasets?
- Is the population spine created by linking admin data sufficiently similar to adjusted 2001 Census records?
- Is it possible to disambiguate variables (for example address) for linked entities?

In summary -

- Work is underway in the UK to try and realise the statistical benefits of the wealth of admin data collected in the UK.
- The proposals put forward by the CIP and IPSS show the potential for statistical analyses afforded by admin data
- Moving from admin data in departmental silos, to collection and presentation of admin data in one place (NeSS), to data linkage.
- There are still obstacles to overcome, including legal obstacles and public opinion.

For more information on the CIP

- CIP website

<http://www.gro.gov.uk/cip>

- To contact CIP please email
populationregister@ons.gov.uk

- Proposals for an Integrated Population
Statistics System

http://www.statistics.gov.uk/downloads/theme_population/ipss.pdf