

---

# Population Statistics Beyond 2011: Research on Census Alternatives

Scottish Privacy Forum  
17 January 2012

Alternative Sources  
National Records of Scotland (NRS)

---

# Overview

1. Background
2. The proposed research
3. How we protect privacy
4. Discussion

---

# Purpose and Scope

- Establish and test alternative models to underpin population and socio-demographic statistics after the 2011 Census
- Feasibility research to report in 2014 (phase 1), followed by a longer term programme of work to pilot, and implement an alternative model(s)
- Under Scottish 2011 Spending Review allocation which also includes wider data linkage framework to support social policy research
- Office for National Statistics (ONS) Beyond 2011 programme for England and Wales established April 2011

---

# Key Drivers

- Cost and viability of traditional census
- User needs and expectations: quality and efficiency
- New opportunities: technology and data
- House of Commons Treasury Committee Report 'Counting the Population', 2008

---

# Key Elements

- Potential building blocks
  - Administrative data
  - Address Register
  - New data collection (census and survey based options)
- Benchmark against 2011 Census
- Use record linkage to learn about the data and develop solutions
- Not all research involves personal data

---

# Record Linkage and the Census

- The traditional census involves a number of data collections and uses identified records and record linkage in the production process
- Census uses: Scottish Longitudinal Study, Scottish Health and Ethnicity Linkage Study
- Organisational and technical infrastructure for the traditional census can be applied
- This project: a wider range of sources set up for other, administrative purposes which can be joined together

---

# Data sources being considered

- NHS Central Register (NHSCR)
- Community Health Index
- Higher Education Student Record
- Further Education Student Record
- School Census
- Electoral Register
- Department of Work and Pensions (DWP)/HM Revenue and Customs (HMRC) Customer Information System
- Northern Ireland No allocations to migrant workers
- Surveys
- Address registers (Census and One Scotland Gazetteer)

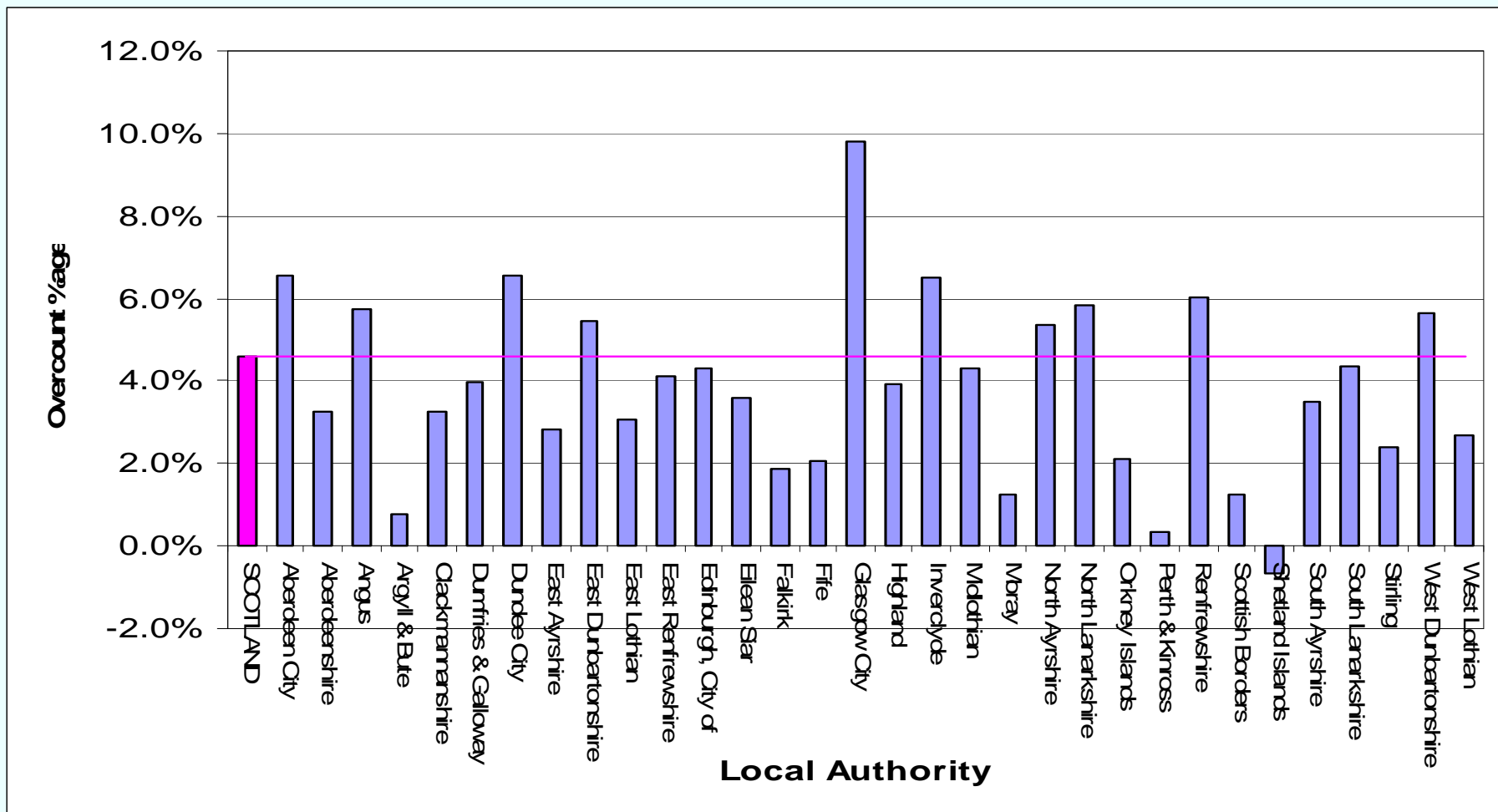
---

# The benefits: why use identified data

- Collected for specific purpose, rules and definitions used not usually consistent with the needs of estimating resident population
- Coverage: geographic or demographic
- Timing and currency: variable
- Quality (completeness, accuracy, consistency): often unknown
- No cross-systems consistency



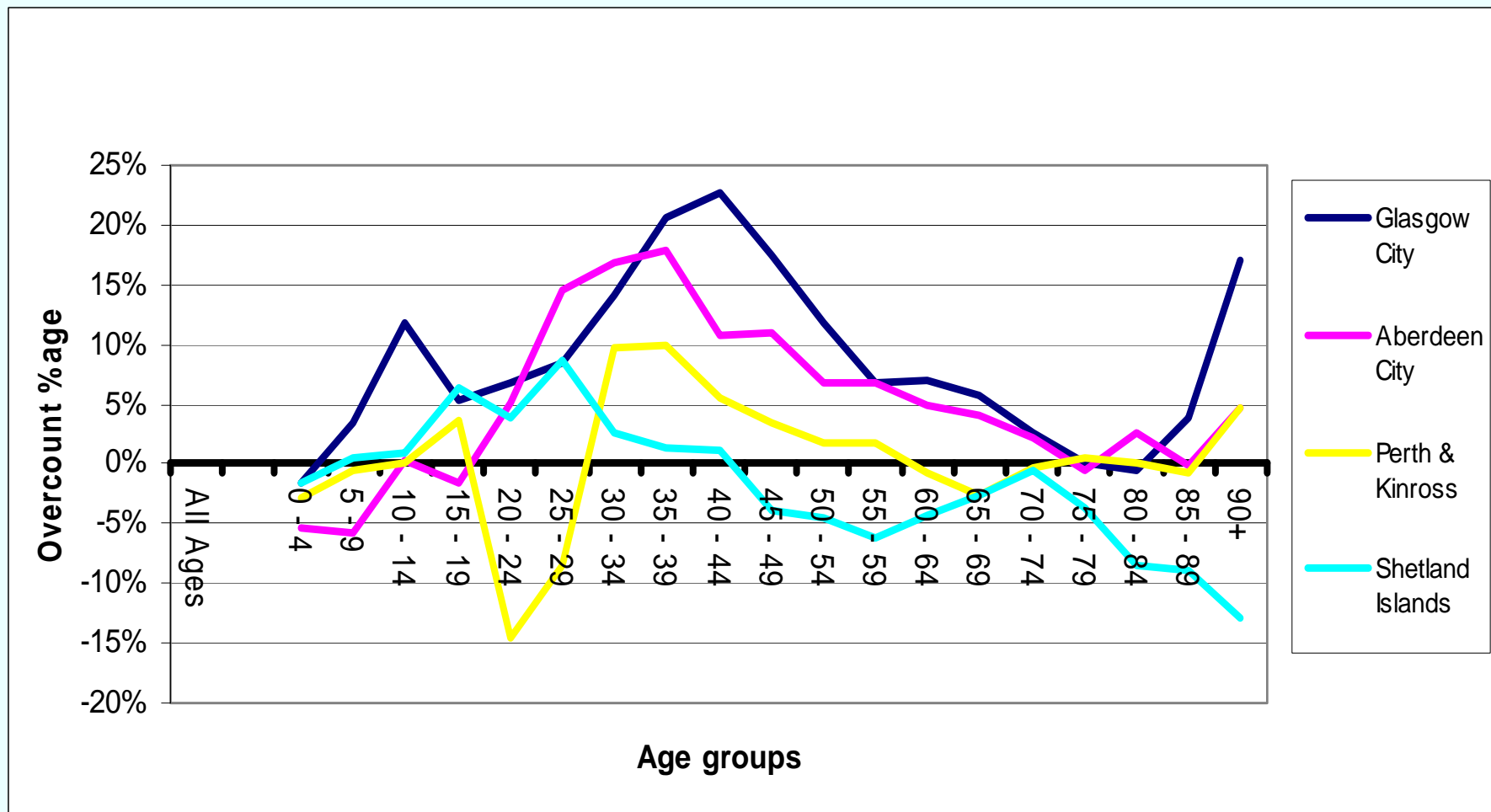
# NHSCR population counts compared with NRS Mid-Year population estimates: 2010, Council areas, % difference



**NATIONAL RECORDS OF SCOTLAND**

preserving the past; recording the present; informing the future

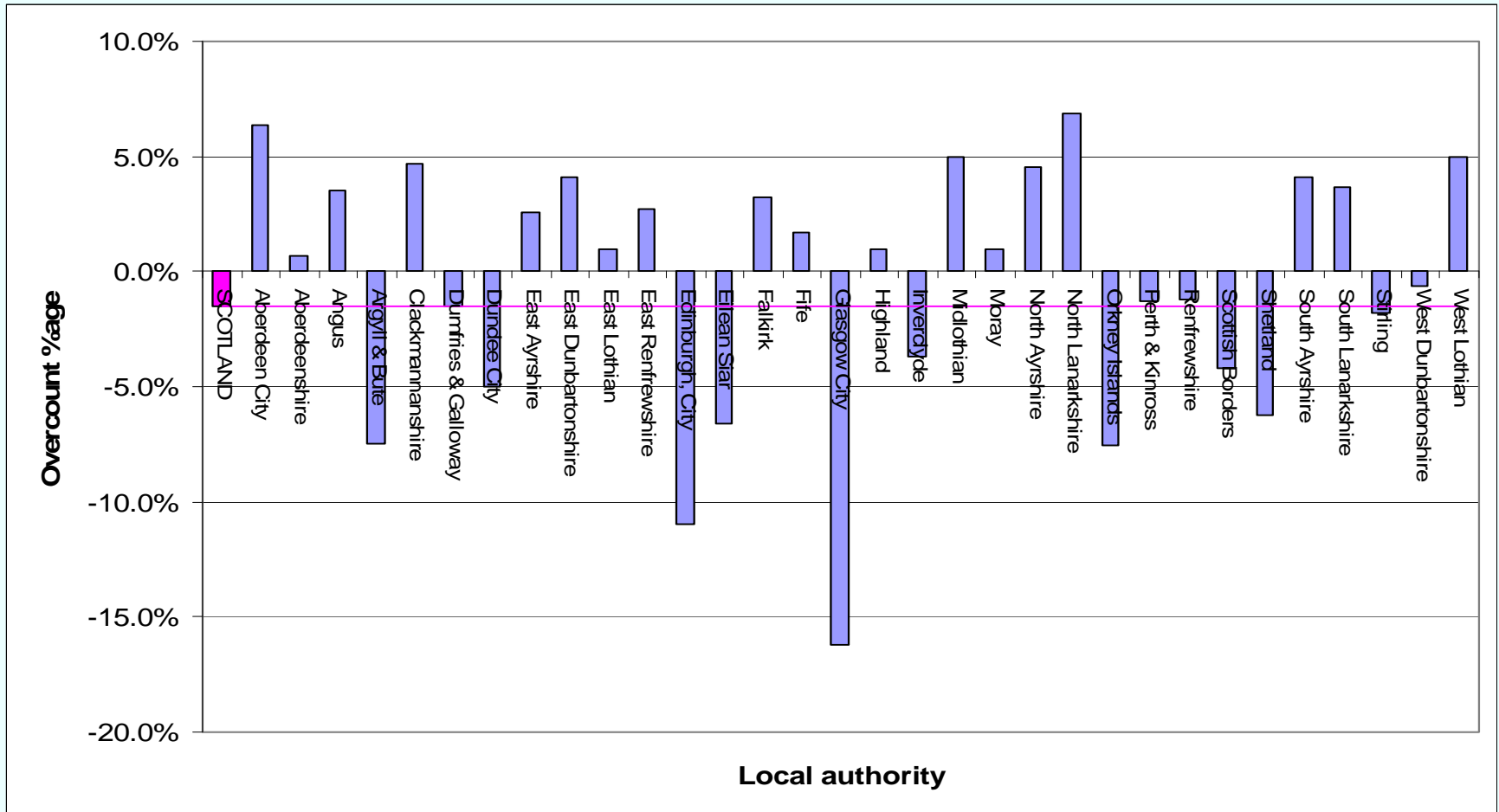
# NHSCR population counts compared with NRS Mid-Year population estimates: 2010, age, % difference



**NATIONAL RECORDS OF SCOTLAND**

preserving the past; recording the present; informing the future

# DWP Customer Information System population counts compared with NRS Mid-Year population estimates: 2010, Council areas, % difference



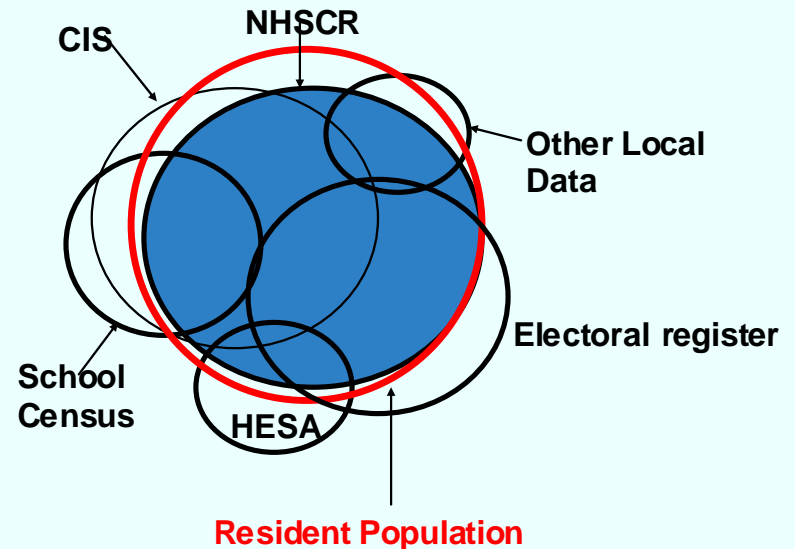
**NATIONAL RECORDS OF SCOTLAND**

preserving the past; recording the present; informing the future

---

# Main conclusions

- No source uniquely capable of providing a solution
- Need to use in combination
- Good understanding of the relevant quality characteristics is a key prerequisite



---

# The research: process and objectives

- Link each source where we have identified data to the 2011 census to understand its coverage and other quality characteristics
- Build a prototype ‘statistical population spine’ [a population count by age, sex, small area] and assess against a benchmark by merging two or more sources
  - Two options: prototype is independent of the 2011 census (may not be possible for a while) or prototype uses census information (more likely in the short term)

---

# Record Linkage

- Bring together records which are likely to belong to the same person
- In the absence of universal person ID number using
  - Name
  - Address (postcode)
  - Date of Birth
  - Genderand a history of name and address where available
- The only way to ensure that people with records in more than one source are not double-counted

---

# Data items: likely (1)

- **NHS Central Register** – name (history), date of birth, postcode of residence, NHS number, registration date, if previously registered elsewhere in the UK, ‘protected’ cases excluded
- **Community Health Index** – NHS number, address
- **Higher Education (HE) Student Record** – name, date of birth, postcode, HE institution, year and type of course, Scottish students registered to study in the rest of the UK
- **Further Education (FE) Student record** – name, date of birth, postcode, FE institution, year and type of course,
- **School Census** – date of birth, postcode, gender, Scottish Candidate Number
- **Electoral Register** – name and address, full register

---

# Data items: likely (2)

- **Surveys** – name, date of birth (age), address, gender
- **Birth Registrations** – name, address, date of birth, mother's name
- **Death Registrations** – name, address, date of birth, date of death
- **Census** – name, address, date of birth, gender, address last year



---

# Data items possible/longer term

- **Higher Education Student Record** – as before for students studying in Scotland
- **School Census** – name
- **Department for Work and Pensions (DWP)/HM Revenues and Custom (HMRC) Customer Information System** – name (history), address (history), date of birth, date of death, marital/relationship status for persons in contact with DWP or HMRC
- Northern Ireland No allocations to migrant workers – address, age and gender
- Other locally held data

---

# Protecting privacy and confidentiality: legal safeguards

- |   |   |
|---|---|
| <ul style="list-style-type: none"><li>• 1920 Census Act</li><li>• Statistics and Registration Service Act 2007</li><li>• Data Protection Act</li><li>• Human Rights Act</li></ul> | <ul style="list-style-type: none"><li>• Confidentiality provisions: unlawful disclosure of personal data is a criminal offence</li><li>• Purpose: processing for statistical purposes only. No feedback to originating administrative systems or use for operational purposes</li></ul> |
|---|---|

---

# Protecting privacy and confidentiality: technical and organisational measures

1. Compliance with Her Majesty's Government (HMG) Security Policy Framework: technical, procedural and physical security controls; independent review of security and Information Assurance process for census 2011 is published on the National Records of Scotland website:  
<http://www.gro-scotland.gov.uk/census/censushm2011/policy-and-methodology/index.html>
2. Separation of matching and content data: analysis on de-identified data
3. Minimum data travel: single transfer to NRS required
4. Personnel: security induction and vetting, follow-up training
5. Access control policies
6. Statistical disclosure control for outputs

---

# Protecting privacy and confidentiality: Public engagement and acceptability

- Legal compliance is not enough, not just about data protection
- Viability of a potential alternative solution or the traditional census itself is dependent on public trust
- Public acceptability research, consultation and transparency of approach; increase awareness and understand concerns as part of the Privacy Impact Assessment (PIA) process

---

# Summary

- Clear public benefit
- Minimum information necessary used and processing tightly specified
- Under strict legal constraints
- Technical and organisational measures to protect confidentiality
- Transparency and Public Awareness

---

# Not just a privacy threat

- ‘Persons, households and firms can live in peace, they are not harassed with unnecessary inquiries. Data security is better because of the reduction in the number of persons handling the data... Only the computer ‘sees’ the data because it is all in electronic format’. Register Based Statistics in the Nordic Countries, UN 2007
- ‘There is evidence that the public is more relaxed about privacy, particularly as a value to be traded off against good or improved public service, than experts and privacy campaigners often realise (cf. Kelsey 2009, Bradwell 2010)’ Kieron O’Hara Review of Privacy and Transparency, 2011

---

Thank you