



Enhancing clinical and public health research capability through record linkage

Professor Carol Dezateux

On behalf of the UK CRC Research and Development
Advisory Group to Connecting for Health

Outline



MRC

Centre of
Epidemiology for
Child Health

- **Rationale**
 - Best Research for Best Health
 - NHS IT
 - Benefits
- **UKCRC RAG to CfH ‘simulations’**
 - Objectives
 - International context
 - Key findings
 - Next steps

Best Research for Best Health

Introducing a new national health research strategy



January 2006



Goal 4: “Manage our Knowledge Resources”



Objectives

- Create a unified knowledge management system to meet the needs of stakeholders
- Use information systems to harmonise and simplify research processes
- Ensure research knowledge is made readily available to professionals in the service, researchers and the public
- Facilitate the application of research outcomes to improve health and delivery of services



The potential of the national IT system



- unique and unrivalled opportunity for research into health
- ‘cradle-to-grave’ electronic patient records with ability to follow the complete patient journey
- demographically, geographically, socially and ethnically diverse large population
- access to coded, structured, longitudinal and comprehensive patient-level data

NHS

Connecting for Health

Better information better health

“A single IT system”



- **unify and simplify the administrative procedures**
- **ensure that procedures and data input are ... undertaken once for multiple uses [‘do once and share’]**
- **populate information systems for the many users and parties interested in health and social care research,**
- **make information supporting regulatory approvals and permissions available to those who need it**



Two key priorities

Support for interventional research

- identify efficiently and comprehensively patients eligible for a specific healthcare intervention (eg therapy or preventative activity)
- facilitate study feasibility assessments and recruitment into trials, and for remote data capture, hence enabling faster and cheaper clinical trials

Support for observational research

- data collected during the course of routine clinical care are used to study the health of the population, the natural history of disease, the safety profile and the clinical and cost effectiveness of healthcare interventions as used in daily clinical practice.



Public engagement and understanding



- “work ... to ensure that everyone using [**?and delivering**] NHS care, including their families and carers, is aware that
 - research is part of the core business of the NHS
 - the quality of NHS care depends on research-based evidence
- anyone using the NHS can expect to be offered opportunities to take part in studies relevant to their needs.”



Why?

The benefits

MRC

Centre of
Epidemiology for
Child Health

Research using electronic records is needed for

- Making the health service safer
- Reliable assessment of different causes of disease
- Identifying effective treatments more rapidly
- Answering public health concerns



Why?

The benefits

MRC

Centre of
Epidemiology for
Child Health

Improved patient management, patient safety and health of the public

- Improves patient management through data feedback
- Allows long term follow up of participants in clinical trials
- Enhances public health surveillance
- Ensures timely responses to public health concerns
- Enables follow up of patients in sensitive settings
- Patient safety
- Public safety



Why?

The benefits

MRC

Centre of
Epidemiology for
Child Health

Improved methodological rigor and data quality

- Validation of data completeness
- Validation of data quality
- Avoidance of bias and misleading conclusions
- Identification of biases in consent and participation
- Avoidance of ecological fallacies



Why?

The benefits

MRC

Centre of
Epidemiology for
Child Health

Efficient large scale research infrastructures with better governance

- Intergenerational studies
- Twin studies
- Infrastructure for large scale evidence [eg UK Biobank, RCTs]

The 'Diamond' Committee



Remit

- to promote collaboration between the UKCRC and Connecting for Health to envisage how the NHS Care Records Service (CRS) can be of benefit to research.

Objectives of simulations

- Inform future development of the NHS Care Records Service (NHS CRS)
- Highlight technical, regulatory and governance issues
- Inform plans for any further simulations and full pilots to test the capacity of the infrastructure, using real patient data with appropriate safeguards when this becomes feasible.

The four simulations

MRC

Centre of
Epidemiology for
Child Health



1. **Interventional Clinical Trial (Rob Thwaites)**
 2. **Surveillance: Pharmacovigilance (John Parkinson)**
 3. **Cohort (prospective) tracking : Biobank (Andy Harris)**
 4. **Observational epidemiology (me)**
- **co-ordinated throughout, with the four ‘Simulation Leads’ meeting and functioning as a team.**
 - **dedicated input from Jeremy Thorpe, Simon Heathfield, Jon Fistein) at CfH**
 - **Liam O’Toole and Ngozi Okwudili-Ince at UKCRC**
 - **4 months!**



International perspective

MRC

Centre of
Epidemiology for
Child Health

Learn from international models

- **Scottish Record Linkage system**
- **Nordic countries**
- **Western Australia**
- **Canada**
- **US**



Lessons from international models



- **Integrated benefits for patient care and research**
- **Leadership**
- **Systematic but federated data architecture**
- **Ownership (scale)**
- **Feedback to contributors**
- **High data quality and completeness**
- **Proportionate governance**
- **Resources for curation, access and analysis**
- **Unique patient identifier throughout**
- **A long haul!**

Existing strengths



MRC

Centre of
Epidemiology for
Child Health

Build on existing strengths in partnership

- Linkage to deaths and births
- Cancer registers
- Validated datasets eg GPRD
- Extensive use of IT in primary care
- NHS Numbers of Babies
- Extensive network of cohort studies and data sources
- Enhanced Hospital Episode System data

Data Sources for Observational Epidemiology



NCRS eg
 Personal Demographics Service
 Personal Spine Information Service
 Transaction Messaging Service
 Secondary Uses System
 Spine Directory Service
 NN4B / Central Issuing System
 Choose & Book, Payment by Results, GP2GP etc
 Electronic Prescriptions

NHS National Collections eg
 Commissioning Datasets
 Mental Health Minimum Data
 QOF/QMAS

Specialist Collections eg
 Cancer/ Diabetes / Renal Audit;
 waiting times; workforce;

Office of National Statistics
 • Birth, Death, ToP, Marriage
The Information Centre
 • Secondary Uses System
 • Censuses & Special Surveys
 eg HSE, NDNS, GHS CEMACH,
 CEPOD, Infant Feeding, etc.....

UNIQUE IDENTIFIER
@ BIRTH / ARRIVAL in UK
NHS NUMBER
CHILD MOTHER FATHER

Primary care

- GPRD, EMIS, THIN et al
- Child health records
- Immunisations [COVER]

Hospital Care Records

- Hospital admissions
- Operative procedures
- A & E / day cases
- HES maternity 'tail
- Special clinics & services
- Fertility (NHS/private)
- Genitourinary medicine
- Intensive Care Networks

Devices/prescribing

- Cochlear implants
- Hip /knee replacement
- MHRA systems

Diagnostics/Imaging

- Ultrasound / Xray [PACS]
- Mammography
- Cytology/Pathology
- Haematology
- Chemical Pathology
- Virology/Microbiology
- Blood Transfusion
- Screening programmes
- HPA surveillance

Registers/databases

- Cancer registers
- Diabetes registers
- Renal registers
- Congenital anomaly registers
- HFEA database
- Cerebral palsy registers
- Down syndrome registers
- Congenital rubella register
- HIV database
- Newborn screening databases
- Central Cardiac Audit
- Juvenile chronic arthritis
- Inflammatory bowel disease
- Dysmorphology database
- Rare disorders
- Public Health Observatories

Cohorts/ Biobanks

- 1946 1958 1970 Millennium
- ALSPAC, ELSA
- MidSpan, Aberdeen, Walker
- Generation Scotland
- UK Biobank
- Newborn Biobanks

Environment

- UK Air Quality archive
- Environmental agency [Landfill]
- Drinking Water Inspectorate
- British Geological Survey
- GIS data [mobile phone masts]
- Superoutput areas / small area microdata

Social Care

- Child Protection, In Care/Adopted
- Elderly care

Income & Benefits

- Benefits, Housing, Income

Education & Employment

- Preschool/day care
- Special Educational Needs
- Pupil Level Annual School Census [PLASC] eg SATS scores
- GCSE, GCE, Higher education
- Occupations and employment



MRC

Centre of
Epidemiology for
Child Health

- **What is in place in the NCRS/SUS.....**

..... and what is not?

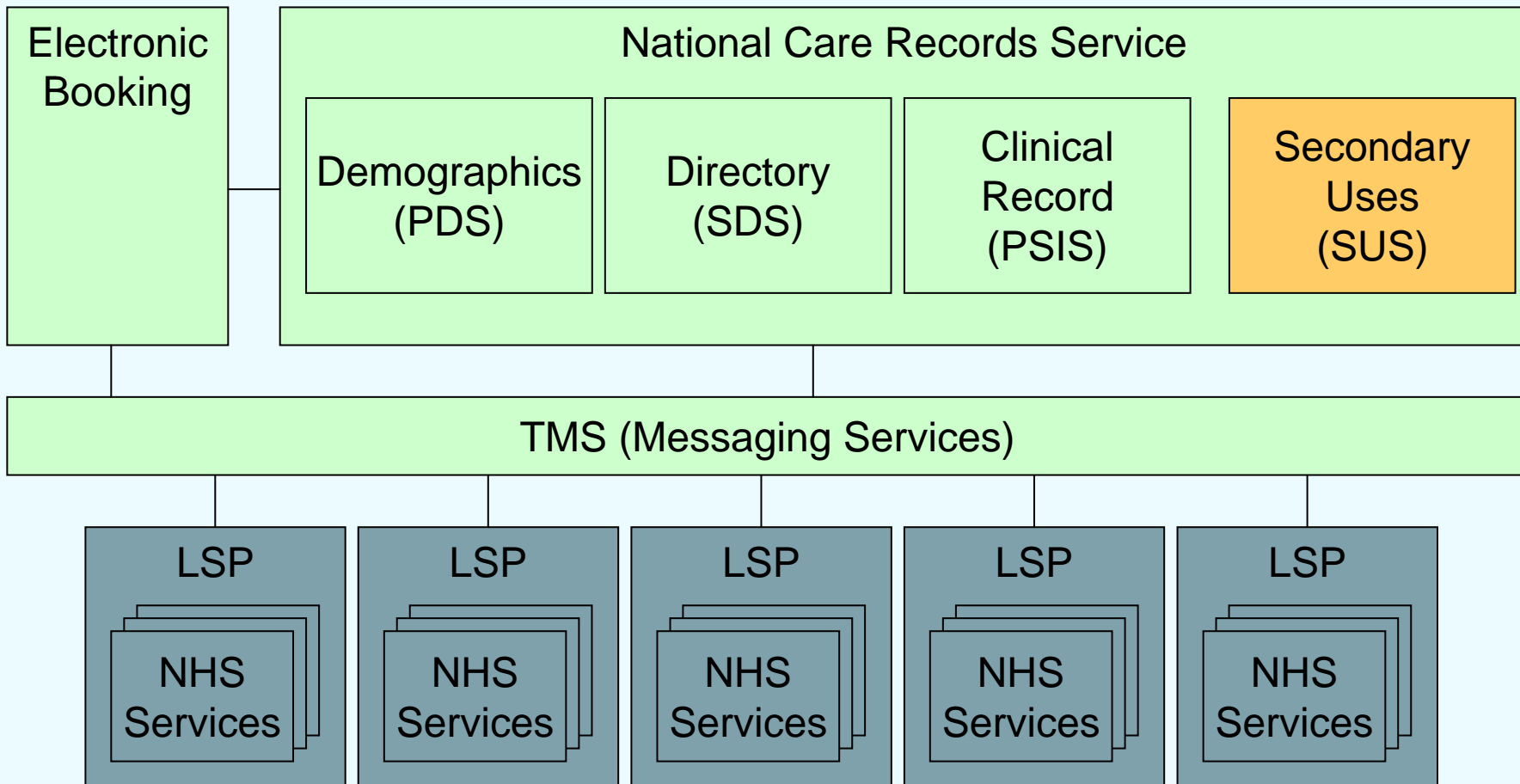


National Programme for IT

MRC

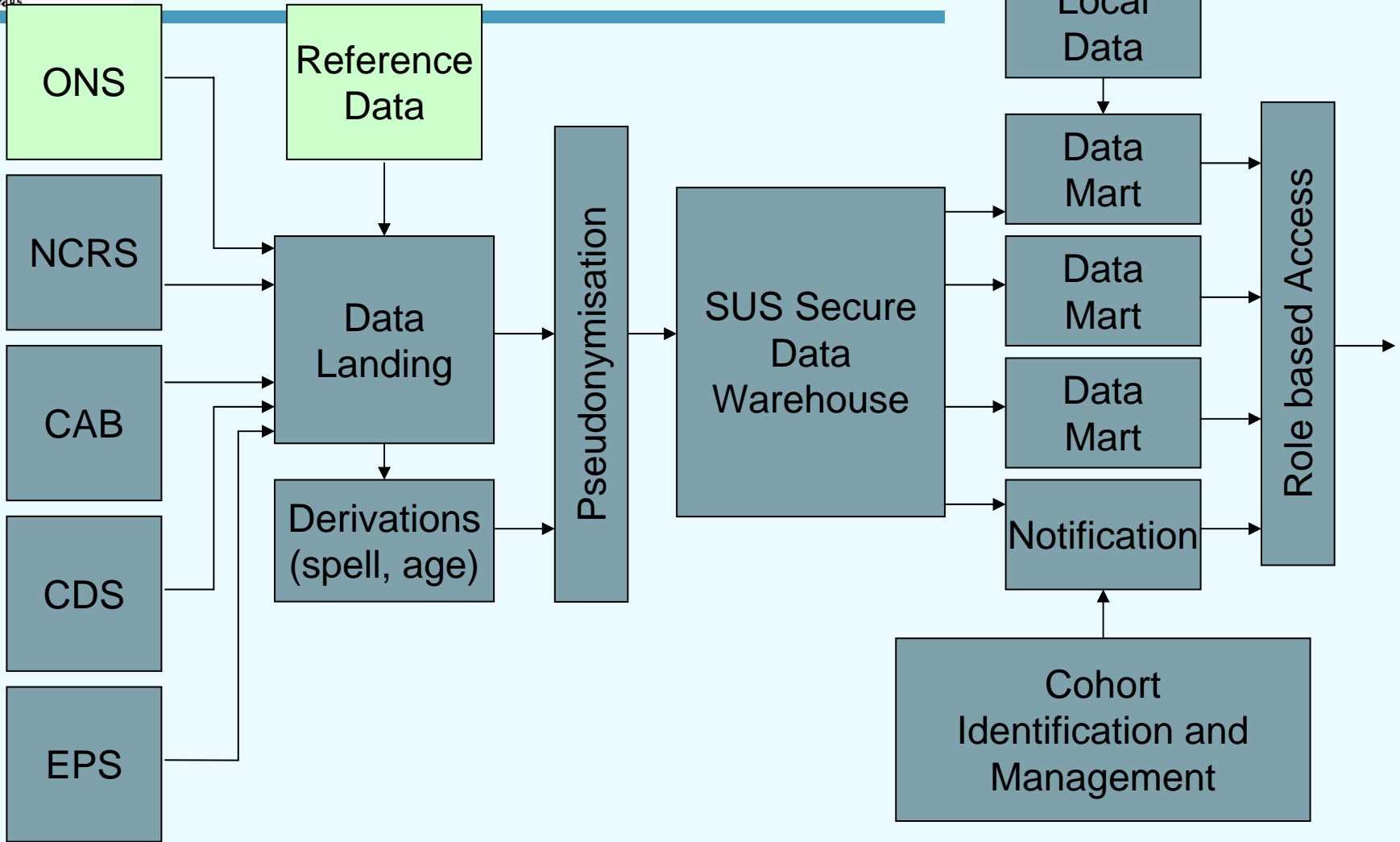
Centre of
Epidemiology for
Child Health

Access to data collected in **operational** care for research, planning...





SUS Architecture





MRC

Centre of
Epidemiology for
Child Health

- **What have we learnt from the simulations?**

Key findings



MRC

Centre of
Epidemiology for
Child Health

Strong consensus across all simulations that

- Research is integral to patient benefit
- Research improves quality of clinical care and data and its assessment
- Leadership is needed
- UK wide solutions
- Individual patient level data needed
- Wide range of data: switchboard rather than warehouse
- Federated data structure supports good governance
- Linkage to primary care and pathology as priority



Key findings: Surveillance/pharmacovigilance

MRC

Centre of
Epidemiology for
Child Health

- **SUS datasets need to evolve to support this function**
- **Full data including text and ‘sealed envelopes’ needed**
- **Legacy (historical) data needed**
- **Way back needed for validation (through ‘honest broker’?)**
- **NHS wide standard incident reporting form for adverse (non drug) events similar to MHRA adverse drug reactions**
- **Data from hospital and day care/out patients needs to be comparable to primary care data (eg GPRD)**
- **Include contracted out NHS services**



Key findings

Clinical Trials

MRC

Centre of
Epidemiology for
Child Health

- **Comprehensive data needed to create recruitment ‘pool’ as well as to assist trial planning and modelling**
- **Complete data across all health care settings and services, including up to date, legacy, linkage at individual level, emphasis on quality**
- **Strict governance but researchers should have access to patient level anonymised data; for recruitment, consented access arrangements will be needed**
- **Standards for recording processing and transferring data needed, validated and customised for trials**
- **Capability for data supply through federated data sources ? SUS or other agency supported by high level strategy**
- **SuS data alone not sufficient**



Key findings

Prospective cohorts: UK Biobank



- **Patient level data including text**
- **Governance and standard processes for data transfer**
- **Mechanisms to update and maintain permissions for research to access demographic and patient record data**
- **Complete medical record including ‘sealed envelope’**



Key findings

Retrospective epidemiology



- **SUS datasets need to evolve to support this function**
- **Full data including text needed**
- **Legacy (historical) data needed**
- **Way back needed for validation (through ‘honest broker’?)**
- **NHS wide standard incident reporting form for adverse (non drug) events similar to MHRA adverse drug reactions**
- **Data from hospital and day care/out patients needs to be comparable to primary care data (eg GPRD)**
- **Include contracted out NHS services**

Recommendations



MRC

Centre of
Epidemiology for
Child Health

Three main areas:

- **Quick wins:**
 - Mandated use of identifiers, communicate relevance of research to NHS and healthcare
- **Short term deliverables**
 - Improve data quality, establish robust governance, develop federated databases, engage stakeholders
- **Development of a UK wide strategy**

Next steps and challenges

MRC

Centre of
Epidemiology for
Child Health



Next steps

- Full report is with Sally Davies and Richard Jeavons but will be in public domain shortly
- Has been included in CSR 2007 spending bid
- High level priority: further work being commissioned

Challenges

- Governance and public understanding
- Infrastructure across NHS and academic communities to support this vision
- Enhancing data quality