MSc in Official Statistics Statistical Computing: PHLS HIV and AIDS Processing System

Andrew Westlake

Survey & Statistical Computing

63 Ridge Road, London N8 9NP, UK

+44 (0) 20 8374 4723

AJW@SaSC.co.uk (E-Mail)

www.SaSC.co.uk

HIV and AIDS reporting section, PHLS

- Redesign of processing and reporting system
- Old system (1994) based on files processed with Epi-Info and Aspect database
- Requirements
  - » Simplify management of data
  - » Improve consistency
  - » Simplify regular reporting
  - » Enable ad hoc reporting

- All done 10 years ago – but was still in use last year

Old System

- **Input Notifications**
  - » HIV Tests (standard form)
  - » AIDS Diagnosis (standard)
  - » ONS Death Reports
- Some notifications on disc, most on paper

- **Data Storage**
  - » Separate HIV and AIDS files
  - » Deaths Added
- **Reporting**
  - » Quarterly standard reports
  - » Ad Hoc inquiries (e.g. PQs)
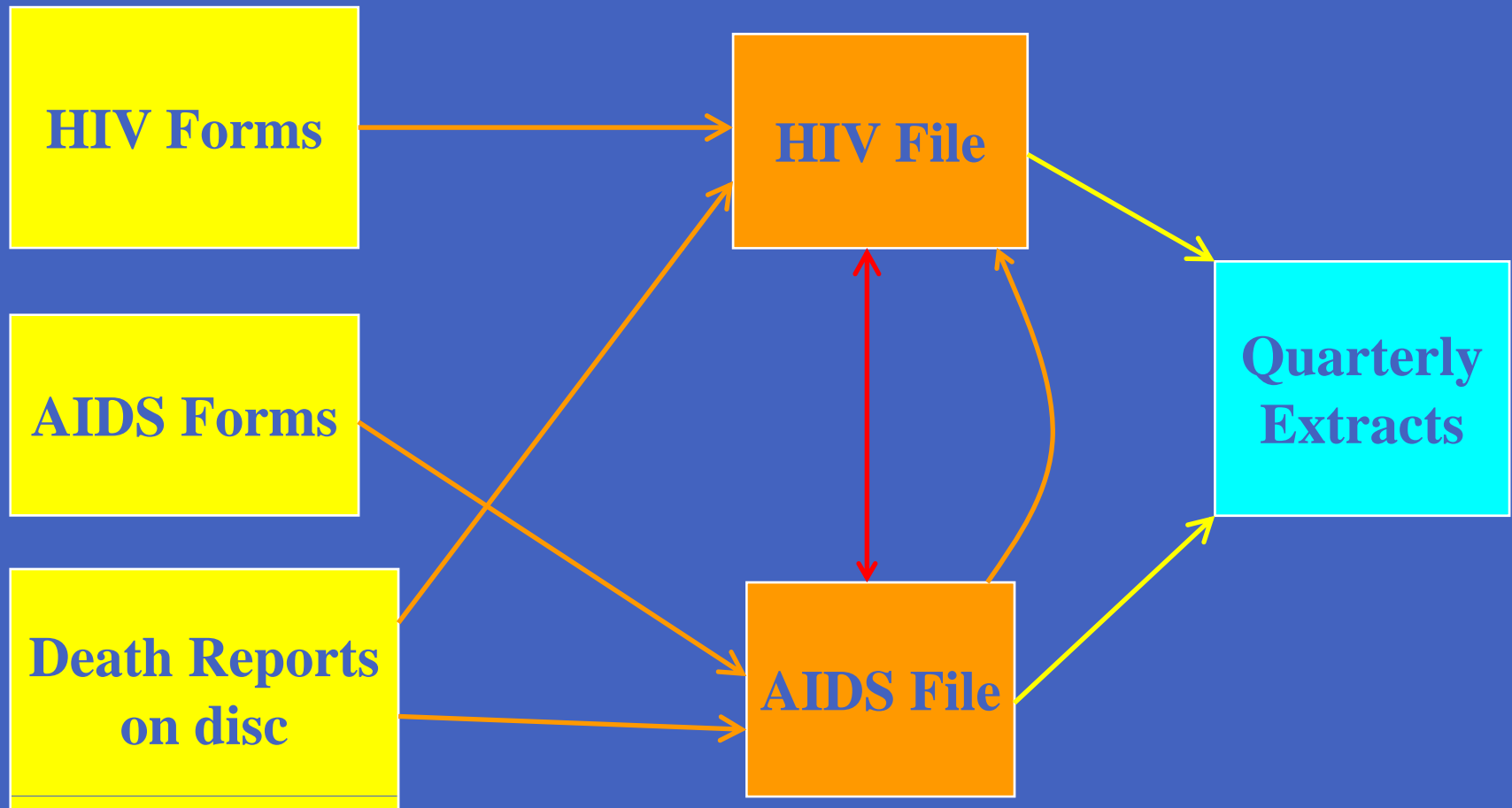 - » Occasional in-depth analysis

- **Processes**
  - » Continuous Data Entry from paper
  - » Regular Linking for reporting extractions
  - » Occasional Matching within and between files

- **Problems**
  - » Duplicates and omissions
  - » Name matching
  - » Repeated linking
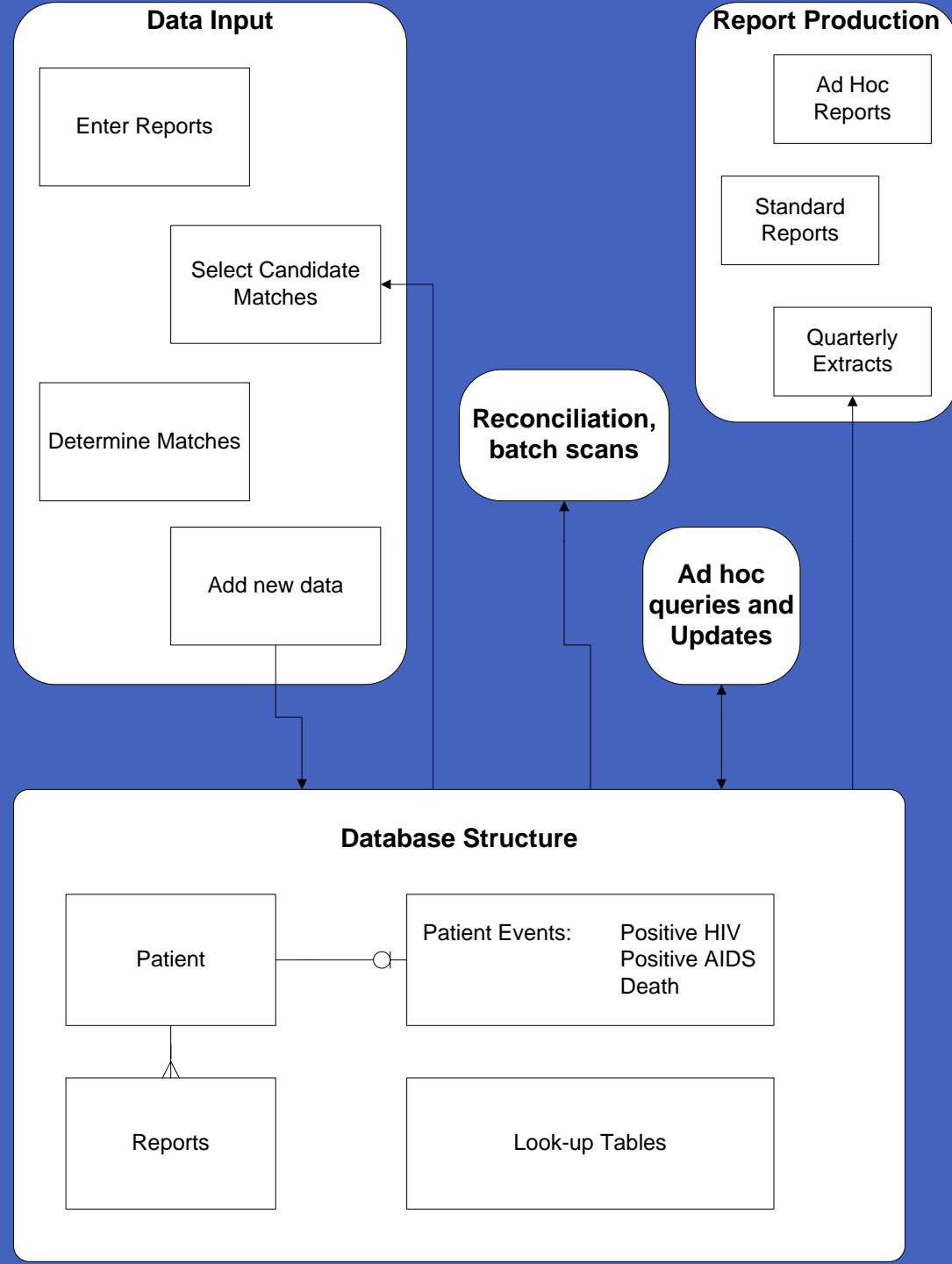
Old Structure

Suggestions for Improvement

- Structure
    - » E-R in Visio
- Processing
    - » DFD in Visio

Initial Thoughts

- Abstraction approach
  - » What are the main components

- Unified database structure
- Focus on Patients
- Do matching once, at initial entry
- Additional components for metadata, auditing
- Extracts to separate reporting system

27-Feb-09

Data Input

Enter Reports

Select Candidate Matches

Determine Matches

Add new data

Report Production

Ad Hoc Reports

Standard Reports

Quarterly Extracts

Reconciliation, batch scans

Ad hoc queries and Updates

Database Structure

Patient

Patient Events: Positive HIV Positive AIDS Death

Reports

Look-up Tables

Full system

Overall Structure

Notification Processing

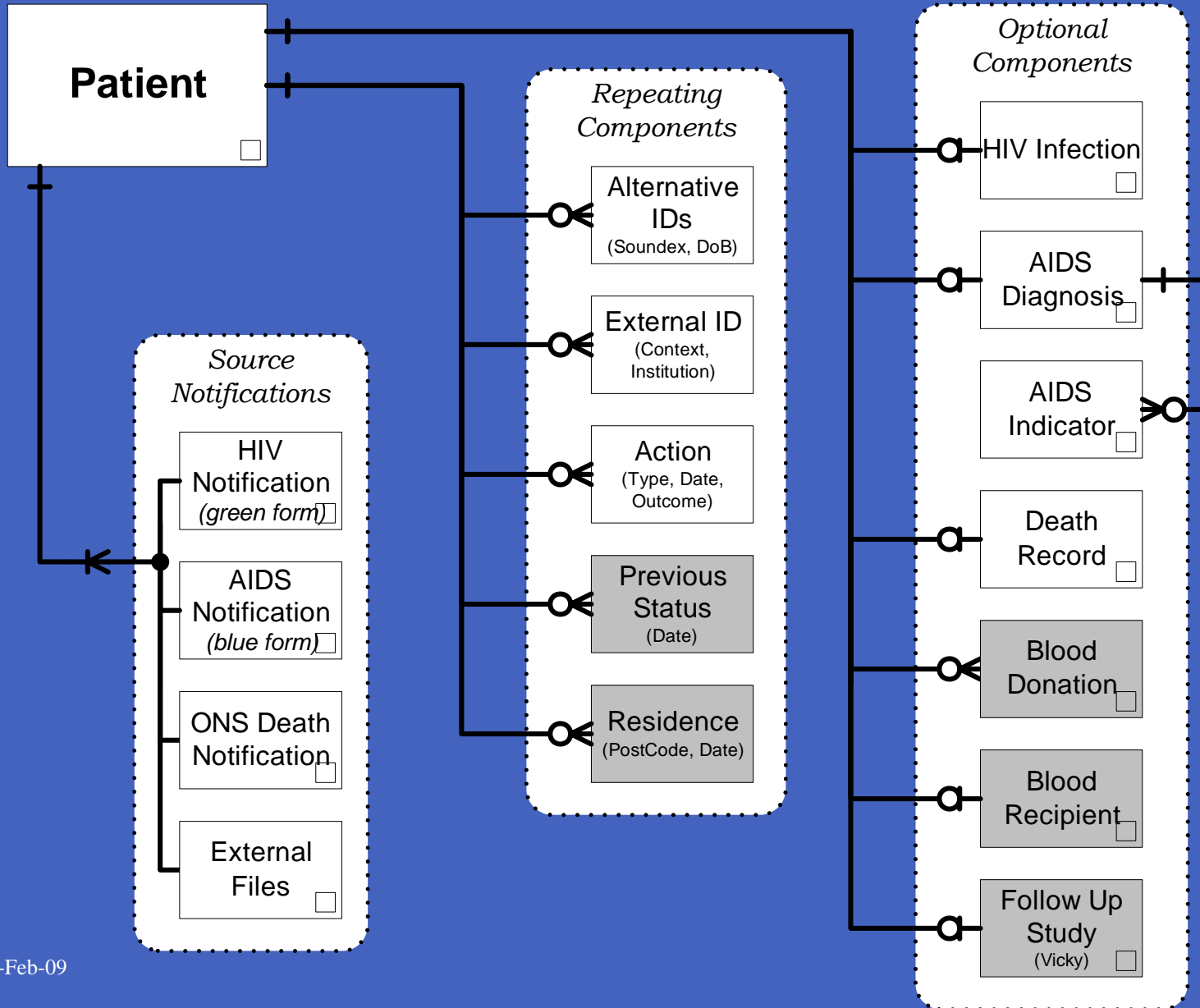**Metadata System**

Patient Data

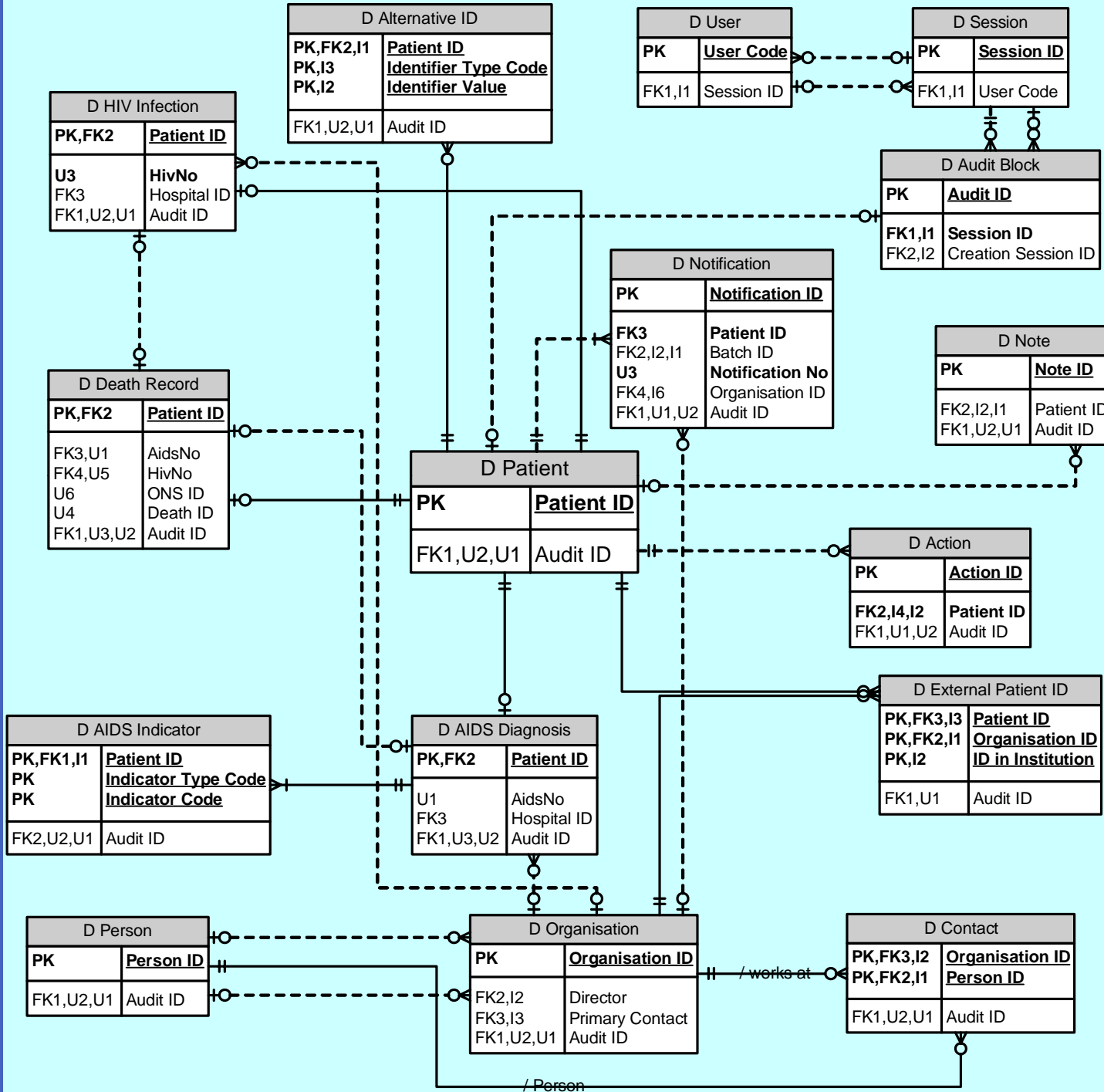**External Organisation**

Reporting Facilities

Auditing System
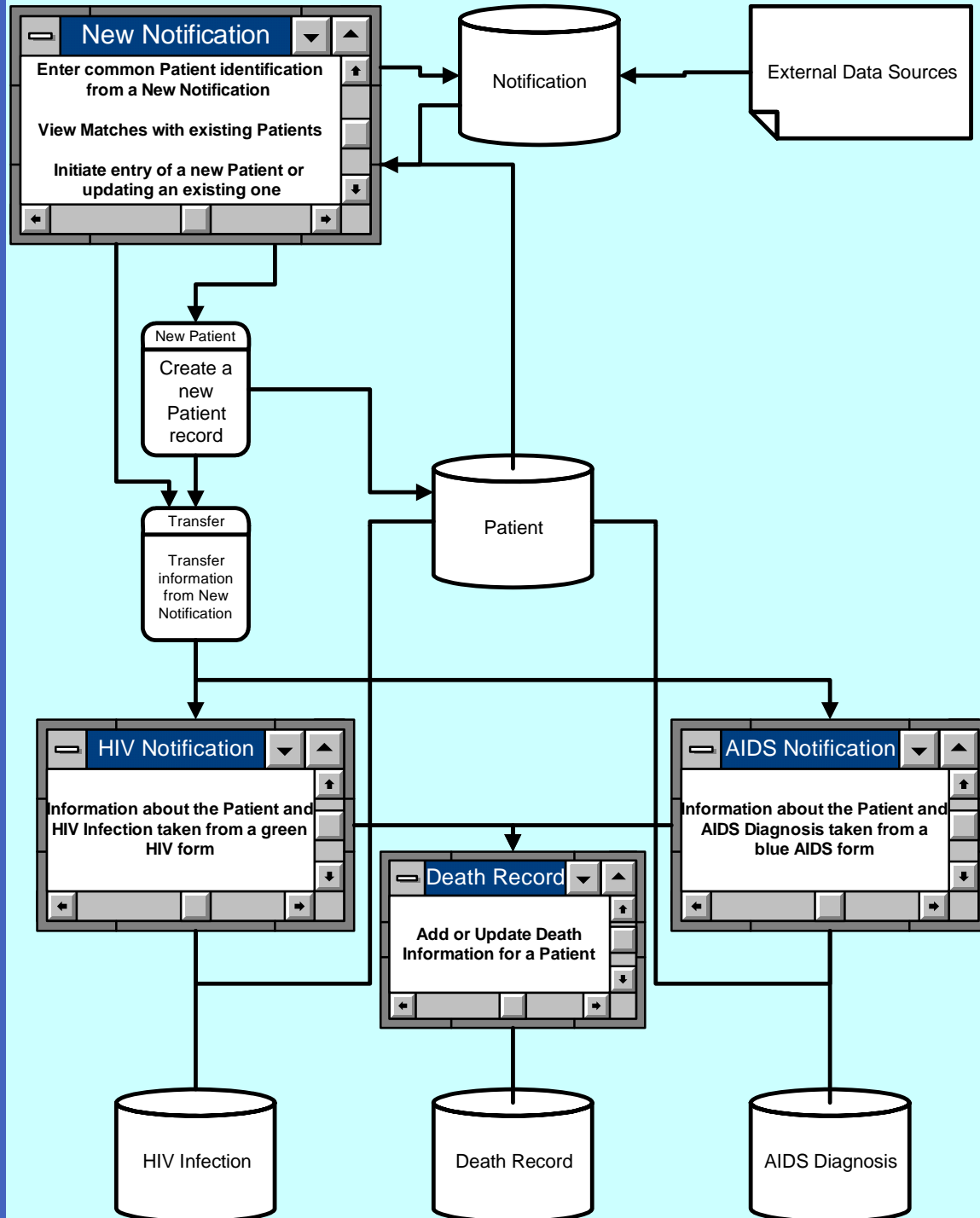
Patients and Notifications

Full HAP Data Structure

Data Entry

- Get Patient details from source

- Look for matching existing patients
  - » May need expert to decide
  - » Create new Patient if needed

- Open appropriate form for source to complete transfer of Patient and details

- Forms know how to handle duplicates (generally take earliest)

27-Feb-09

Extraction for Quarterly Reporting

- Separate Extraction Database template
  - » Linked to main database
  - » Set of associated Excel spreadsheet templates
- New copy each quarter
  - » Populate by Extraction of information
  - » Snapshot of Patient status – available for re-analysis
- Set of aggregation queries
  - » Copy results into spreadsheets as source
- Reorganise with Pivot tables
  - » Create publication versions

Summary

- Old system based on data sources
    - » No analysis of usage requirements
- Need to think about underlying entities and processes
- Initial removal of redundancy makes later processing more efficient
- Data store continually updated
    - » Need snapshots for repeatable analysis

- What could be different now?