Activities and Findings:
- We addressed and solved some engineering challenges around ontology creation, data access and data integration, using tested methods to map datasets to ontologies.
- We used real-world patient datasets to test our solutions, which necessitated the assessment and management of data quality; and
- We confirmed that automated identification of diabetes patients can be specified systematically as a solution supported by semantic retrieval.

Discussion:
- Primary and integrated care “big data” tasks that require automating include semantic integration of clinical data from multiple EHRs; assessment and management of the provenance and quality of EHR data such as reasons for visit, chronic conditions and diagnoses, pathology tests and prescriptions; and preservation of meaning of the data, information and knowledge as they may be perceived and interpreted by clinicians and researchers.