American Medical Informatics Association

Leaders Shaping the Future of Biomedical and Health Informatics

SECONDARY USES AND RE-USES OF HEALTHCARE DATA: TAXONOMY FOR POLICY FORMULATION AND PLANNING

This taxonomy is to be used as a resource in developing plans and policies related to secondary uses of healthcare data. The first section is meant to be inclusive of all categories or classes of secondary uses of healthcare data (in **bold**), but not exhaustive in enumerating the entire set of detailed uses in any given category. In contemplating a plan or policy relative to secondary uses of data, one should be able to refer to the taxonomy and consider how the plan or policy might apply to each class or category of use listed in the taxonomy.

The second section of the document lists factors that may be considered in formulating plans or policies for the secondary uses and re-uses listed in section one. For example, policies for re-use of data might be more restrictive for identifiable data or for data obtained from a vulnerable population.

The third section of the document lists the requirements and restrictions that might be imposed on the secondary use of data on the basis of the uses and factors for consideration. One would then be able to relate policy statements to a particular secondary use of data. For example, if a policy stated that "All secondary use of data to identify markets and promote sales requires prior patient consent" one would know that consent should be obtained from all patients contributing to the database before their data could be shared with manufacturing and sales firms that would target them as potential customers.

The final section of the document lists existing or potential sources of secondary data.

1) **Secondary uses and re-uses of data**
   - a) **Protect and enhance public health**
     - i) Enable and support biosurveillance
       - (1) Monitor and report vital statistics
       - (2) Monitor and report biometric demographics (e.g. weight, height, blood pressure, normal lab values)
       - (3) Identify, monitor, and report health and illness trends
       - (4) Identify, monitor, and report infectious diseases (e.g. culture, serology, DNA/RNA probe results)
     - ii) Export data to health registries
       - (1) Cancer or rare disease registries
       - (2) Drug and device registries
     - iii) Report toxic exposures (e.g. smoking, Agent Orange)
   - b) **Develop security and confidentiality algorithms and test de-identification routines**

- **c) Conduct research**
- **d) Create and maintain terminology and representation formalisms**
- **e) Develop and apply decision support for health care providers**
  - (i) Develop and test the efficacy of decision support algorithms
  - (ii) Develop order sets, rules, and alerts
- **f) Support quality of patient care**
  - (i) Manage quality and outcomes
  - (ii) Manage staffing and resources
  - (iii) Develop and assess quality indicators
  - (iv)Support quality reporting (e.g. HEDIS)
- **g) Improve patient safety**
  - (i) Conduct pharmacovigilance (post market drug and device surveillance)
    - (1) Detect and analyze adverse and sentinel events
    - (2) Support risk profiling
  - (ii) Monitor and survey to prevent patient adverse events
- **h) Manage personal health**
  - (i) Provide patient-specific feedback and assessments of progress toward health goals
  - (ii) Maintain personal health records
  - (iii) Provide links to knowledge resources based on personal health information
- **i) Educate and credential healthcare providers and assess training activities** (e.g. types and outcomes of procedures)
- **j) Analyze and Manage Finances**
  - (i) Conduct automated billing, claims processing
  - (ii) Analyze activity-based charge capture, cost accounting
  - (iii) Develop predictive models of costs and accounting
- **k) Detect fraud and illicit activity**
  - (i) Detect illegal and inappropriate activity (e.g., Medicare upcoding)
  - (ii) Report drug screen results to detect illegal drug use
- **l) Identify markets and promote sales**
  - (i) Conduct market research
  - (ii) Target marketing to physicians
  - (iii) Target marketing to patients and families

2) <u>Factors influencing authorization for secondary use of healthcare data</u>

- **a) Identification Status**
  - i) Patient-identifiable data
  - ii) De-identified data (HIPAA definition)
  - iii) Anonymized data
    - (1) No linkage possible (alteration of PHI, precluding linkage)
    - (2) Relinkable data
    - (3) Linked with protected key (trusted third party)
- **b) Consent provided at the time of data collection**
  - i) No consent by the individual
  - ii) Consent by the individual
    - (1) Broad and unspecified

(2) Time-limited consent
(3) Consented for partial, source specific use (e.g., no psychiatric data)
(4) Consented for the particular type of secondary use

c) Demographic representation
i) Age
ii) Race
iii) Gender
iv) SES
v) Insurance status

d) Focus on a vulnerable population (e.g. prisoners, pregnant women, undocumented immigrants)

e) Original collector and aggregator of the data
i) Government
ii) Health Plan
iii) Other private entity

f) Proposed secondary user of the data
i) Government agency
ii) Academic institution
iii) Private, not-for-profit entity
iv) Private, for-profit entity

g) Funding source for secondary use
i) Government agency
ii) Academic institution
iii) Private, not-for-profit entity
iv) Private, for-profit entity

h) Financial compensation to data collector or data steward for providing data to a second party
i) No compensation
ii) Compensation

i) Beneficiary of secondary use
i) Society
ii) Researcher
iii) Academic institution/medical center
iv) Private, for-profit entity (e.g., financial gain)

j) Disclosure of secondary use
i) Not disclosed publicly
ii) Publicly disclosed
(1) Disclosure of results only
(2) Disclosure of research methods utilized
(3) Disclosure of analytic principles that guided the use of the data

3) <u>Requirements imposed on secondary use of healthcare data</u>

a) Required level of consent and authorization
i) IRB evaluation not required
ii) IRB evaluation required

(1) No consent by the individual required
(2) Consent by the individual required

b) Compensation of patients

i) No compensation required
ii) Compensation of individual patients required

4) <u>Existing and potential sources of data for secondary use</u>

a) Public Use Datasets

i) Medicare
ii) Medicaid
iii) CDC surveys (some Primary data use, e.g. NHANES)

b) Private Datasets

i) Open-source data
ii) Commercial use datasets (at patient level)
(1) Pharmacy benefit/claims manager
(2) Provider databases
(a) Individual providers
(b) Aggregated data from provider consortia
iii) Consortium databases
(1) caBIG
(2) CTSA recipients
(3) University Health Systems Consortium
iv) Aggregated clinical repositories hosted by HIT vendors
v) Personal health records, including patient-entered data
vi) Health Information Exchanges (RHIOs, etc)