MITRE SHR\_Completeness\_v01

Rough Notes

23.Jun.16

**INTRO**

Documentation on determining differing levels of completeness in a patient’s Standard Health Record. Prioritizing and ranking data elements in the Standard Health Record based on importance and use.

**CONSIDERATIONS**

* Primary object of interest is the individual patient
* The prioritization and different levels should balance relevancy to individual patient vs population health/clinical research applications
* Difference between complete info about patient vs complete info about patient encounters. Patient w/no care encounters and empty record has a ‘complete’ record with respect to the health care process, but a blank one with respect to the patient
* Completeness at different granularities. Synthesis of record as a whole vs modular components of record (some data elements have incredible detail of info, while others have just one)
* Defining completeness of record with respect to content vs context of use. Intrinsic vs extrinsic = content vs use. Is the record ‘complete’ for emergency care scenarios but not for continuity of care scenarios? Scenarios = longitudinal patient record/primary care, population health researcher, emergency care scenario.
* Completeness of data is dependent on definition of completeness being used
* Is data elements quality a combination of breadth, depth, scope? How do you balance those factors?
* Identify quality of each data elements before identifying different levels of completeness
* Must be explicit in choice of completeness definition
* Medications have highest predictive correlation with diagnostic codes, sensitivity (completeness) of data generally went medication → consultation/referral → lifestyle smoking/alcohol → socioeconomic data [4]

**METHODS**

Method for determining completeness of the health record:

1. Record begins with all value sets as “null” (placeholder for data that is not known)
2. Determine what data element the health data point belongs under
3. Data point is entered. At this point it is not “null”. It is either a zero, a quantitative value (numeric value, ratios, etc), or qualitative entry (nominal, ordinal, interval values)
4. Determine what time/date the health data point was entered (metadata)
5. Determine who is collecting that data point (metadata. Can be individual, or multiple people. MD, nurse, clerk, receptionist, patient, caregiver, etc)
6. Determine if that data is subject to change (metadata. ex: medical history data won’t ever change)
7. Determine if the data was “certified” on entry (metadata. was there another person who took a second look at the data point, if so, who)
8. Determine if the data was validated by an independent entry by another person (metadata. or was it carried/copied forward at a new setting of care)
9. Collect next health data point. Determine if the data point is a reconfirmation (metadata. ex: updating medication as patient changes site of care). If not, repeat steps 2-8 for each health data point collected
10. Determine the top ~6 or so possible use case scenarios of patient health data
11. Prioritize the use cases by frequency of occurrence
12. Determine broad ‘levels’ of completeness based on full completeness of data elements per case scenarios (1 level is complete for only 1 scenario, next level complete for half of the scenarios, etc)
13. For depth of data, use EHR’s completeness model as reference (4 different types of definitions) and create data scenarios with varying levels of depth

Possible method for determining health data quality:

1. Use AHIMA’s data quality model as a checklist (10 factors) for each data value in SHR v01

**RESEARCH NOTES**

Data integrity risks in data collection workflow [6]:

* Organizations using voice recognition without a validation step in place
* Ensure process where providers review, edit, approve dictated information
* Cloned documentation creating redundancies
* Automated insertion of previous/outdated information through EHR tools, when not modified to be patient-specific/pertinent to the visit
  + Information copy pasted from different patients record
  + Vital signs that dont change from visit to visit
* Documentation template does not exist for a specific problem/visit type
* Templates designed to meet reimbursement criteria encourage over-documentation to meet reimbursement requirements even when not medically necessary
* Patient identification errors = wrong information documented for wrong patient
* Some EHRs dont have functionality for multiple providers to sign a particular data contribution. Metadata issue

Determining health data quality [2]:

* Data accuracy: extent to which the data is free of error
  + Consistent integer format for lab values
  + Correct, unambiguous data elements labels - dosages, drug interactions, etc
  + Less chance of inserting wrong information. I.e. plugging in 5 inches for height instead of 50 because data elements name was misinterpreted
* Data accessibility: level of ease and efficiency at which data can be obtained within a well protected/controlled environment
  + Is the data legal to collect? Other boundaries, constraints? Privacy issues?
  + Are there established guidelines for data ownership?
* Data comprehensiveness: extent to which data within the entire scope are collected
  + How the data will be used is clarified. This is different from defining the rationale for data use. This is more a focus on the process of data use rather than purpose.
  + Identify end users of data
  + Cost-benefit or impact study for when collected data is increased is known
* Data consistency: extent to which data is reliable, identical, reproducible by different users across applications
  + Value of data is same across systems/applications
  + Minimized conflict between related data items/values. Ex: drug dosing
  + Is the formula for determining some values correct/validated? Ex: calculation for BMI is based on comprehensive evidence?
* Data currency: extent to which data definitions & algorithms are up to date.
  + Is the algorithm used to calculate the value set up-to-date + consistent with evidence?
  + Is the data elements definition name + value set labels consistent/up-to-date with current standards?
* Data definition: specific meaning of healthcare-related data element
  + Clear documentation of data name, definitions, attributes so that data is not misinterpreted. Ex: there’s a distinction between race and ethnicity
* Data granularity: level of detail at which attributes and characteristics are defined
  + Data is collected at appropriate level of detail/granularity
  + How granular? Statistics can be collected at daily, weekly, monthly intervals to see trends/long range planning
* Data precision: degree to which measures support their purpose, and/or closeness of two or more measures to each other
  + Precision is how close the values are with repeated measurements. Ex: weigh something 5 times in a row, those 5 values are the same = high precision
  + Defined acceptable values. Ex: gender can be male, female, unknown
  + Well defined process of data collection. Ex: collecting BP data before anti-hypertension medication vs after anti-hypertension medication has different purpose. This process + purpose should be defined
* Data relevancy: extent to which healthcare-related data are useful for the purposes which they were collected
  + The rationale for collecting the data element is clarified
  + Ability to link the data to other data for comparison - establishing relevancy
* Data timelines: availability of up-to-date data within the useful, operative, indicated time
  + Use of data within a specific timeframe. Ex: vitals should be taken once per visit for regular consultation, but every 15 mins for a critical patient

Data quality assessment [5]:

* For each data element, determine values representing good/bad quality data (how much data is needed to make it a data element with a high quality of recorded data)?
* Consistency: can the data be matched across data stores?
* Timeliness: degree to which data represents reality from the required point in time. Accuracy decays with time (would % completeness decrease over time for a health record because of increased risk of inaccurate untimely information?)
* Validity: data conforms to the syntax (format, type, range) of its definition. In health records case, it would be following SNOMED CT/HL7 standards
* Accuracy: how closely does the data describe the real world object?
* Each data element can have a % of completeness

Determining data elements weighting [5]:

* Which data elements are critical to the use case?

Use cases [7]:

* Physician office visits: 928.6 mil (2012)
* Emergency department visits: 136.3 mil (2011)
  + Injury related: 40.2 mil
* Outpatient visits: 125.7 mil (2011)
* Inpatient care: 35.1 mil (2010)

Standards for the clinical structure and content of patient records to rank SHR data elements [1]:

Scenarios where health data can be used:

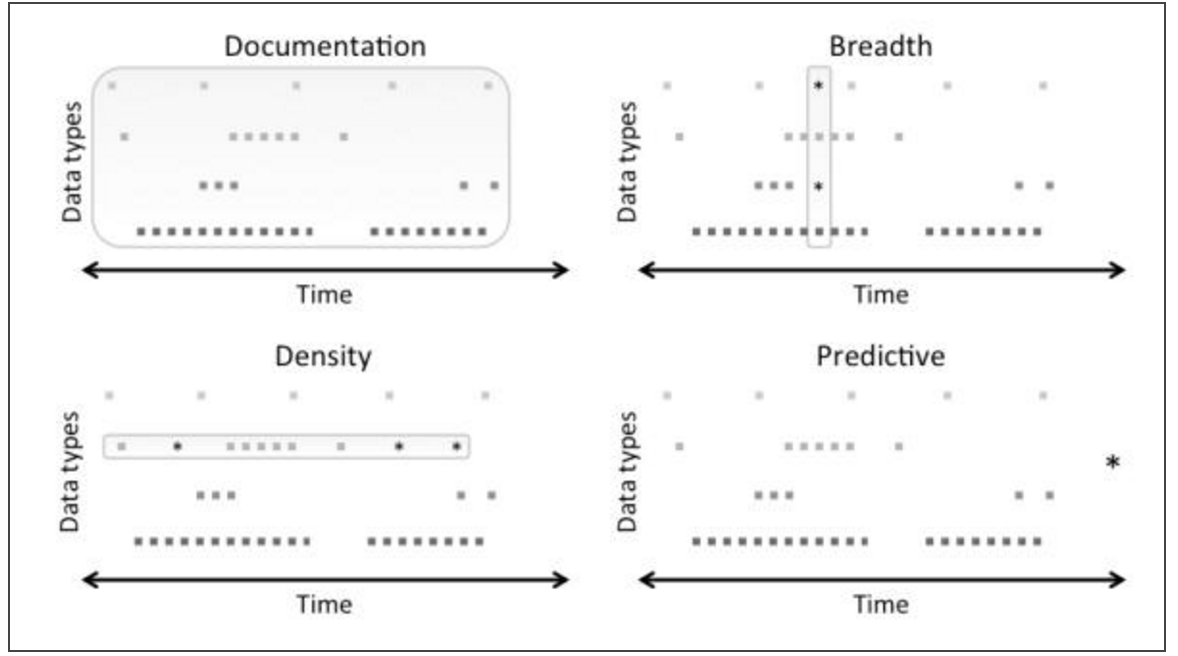
* Admission: hospital admission record
* Handover: handover of patient care from one care provider to another
* Discharge: clinical information recorded in discharge record
* Outpatient: clinical information recorded in outpatient setting
* Referral: clinical information in referral communications between care providers
* Core: priority for inclusion in EHRs. priority based on priority coding defined within SNOMED CT (standardized multilingual vocab of clinical terminology for exchange of clinical health information)

Overlap in data elements between Academy of Medical Royal Colleges and SHR v01 (bolded means overlap exists. Nesting below bolded data elements are possible value sets) + use across types of records. Scenarios are C=core record, A=admission record, H=handover record, D=discharge record, O=outpatient record, R=referral record. Use in more scenarios = greater weight given to the data element?:

* **Patient Identification**
  + **patient name(s) (C A H D O R)** 
    - **First + last name**
    - **Patient preferred name**
  + **date of birth (C A H D O R)** 
    - **Day, month, year**
  + **administrative sex (C A H D O R)** 
    - **Sex at birth**
  + **address(es) (C A H D O R)** 
    - **Place of residence**
  + **telephone number(s) (C A H D O R)** 
    - **Mobile, work, home**
  + **email address(es)** **(C A H D O R)**
  + **emergency contact(s) (C A H D O R)** 
    - **First + last name**
    - **Relationship**
    - **Telephone number**
  + **legal guardian (if minor) (C A H D O R)**
  + **preferred language (C A H D O R)**
* Patient Support
  + payment source
  + insurance identifier(s)
  + **care team members (C A H D O R)** 
    - **First + last name**
    - **Relationship**
    - **Telephone number**
  + **consent for data sharing (C A H D O R)**
  + ***healthcare proxy* (C A H D O R)**
    - **First + last name**
    - **Contact details**
  + ***advance directives* (C A H D O R)**
    - **Existence of document**
    - **Location of document**
    - **Copy of document**
  + ***health goals* (C A H D O R)**
  + *preferred pharmacy*
* Current Health and Care Plan
  + **current medications** 
    - **Name (C A D O R)**
    - **Form (C A D O R)**
    - **Route (C A D O R)**
    - **Dose (C A D O R)**
    - **Frequency (C A D O R)**
    - **Additional instructions (C A D O R)**
    - **Do not discontinue warning (C A D O R)**
    - **Reason for medication (C A D O R)**
    - **Medication recommendations (C D O R)**
    - **Medication status (C R)**
    - **Medication change (C A D O R)**
    - **Reason for medication change (C A D O R)**
    - **Medicine administered (C O )**
    - **Reason for non administration (C O )**
    - **Relevant previous medications (C A O R)**
  + **allergies (drug, food, environ) (C A H D O R)**
    - **Causative agent (C A H D O R)**
    - **Description of rxn (C A H D O R)**
    - **Probability of recurrence (C A H D O R)**
    - **Date first experienced (C A H D O R)**
  + **problems** **(C A H O )**
  + contraindications and intolerances
  + vital signs
  + blood type
  + **medical appliances or devices (C A D O R)**
  + *lab tests and results (recent)*
  + ***health concerns* (C A O R)**
  + ***plan of treatment (care plan)*** **( R)**
  + ***Disabilities* ( A H D O R)**
* Health History
  + past hospitalizations (past year)
  + **past major procedures (C D O )**
  + immunizations
  + ***past outpatient encounters (past year)* ( O )** 
    - **Contact type ( O )**
    - **Consultation method ( O )**
    - **Purpose of contact ( O )**
    - **Appointment time ( O )**
    - **Time patient seen ( O )**
    - **Time consultation finished ( O )**
    - **Specialty ( A H O )**
    - **Service ( A H O )**
    - **Responsible healthcare pro. ( O )**
    - **Care professionals present ( O )**
  + ***personal health history* ( A H O R)**
  + ***family health history* (C A O R)**
* Social/Environmental Factors
  + **race/ethnic group ( A D O R)**
  + **religion** **( A D O R)**
  + national origin
  + **gender identity ( A H D O R)**
    - **Male, female…**
  + marital status
  + *income level*
  + ***education/literacy level* ( A H D O R)**
  + *food status/security*
  + *transportation availability*
  + *health services accessibility*
  + ***housing situation/security* ( A D O R)**
  + ***employment status/security* ( A D O R)**
  + *stress factors*
  + ***social isolation/exclusion* ( A D O R)**
  + ***domestic violence or abuse* ( A D O R)**
  + ***risks to patient, provider, or third party* (C A H D O R)**
* Behavior
  + ***level of physical activity*  ( A O R)**
  + **smoking status ( A O R)**
  + ***drug and alcohol use* ( A O R)**
  + ***medication compliance* ( A H O R)**
* *(standard health record data to add later)*

EHR completeness model [3]:

This is 4 different approaches to defining a complete patient record (based on literature review paper) all % based on data from NY presbyterian hospital patient records (~2mil patient records):



* Documentation: all observations made during a clinical encounter are recorded. It’s a measurement of fidelity of the medical documentation process
  + Rate of completeness based on patient records since (late 80s? I believe): 99% inpatient visits, 74% outpatient visits, 95% emergency visits
* Breadth: secondary use scenarios like quality of care assessments, patient cohort identification, clinical research uses, need a broader set of data. More complete may mean a breadth of data available. Good starting point: looking at lab results, medication orders, diagnoses, sex, date of birth
  + Rate of completeness based on breadth starting point: 10%
* Density: frequency and interval of data points over time. Enough sufficient numbers and data points. Useful for scenarios where algorithms use that database of information to determine clinical trial eligibility, clinical research use
  + 28% had at least 5 recorded events with admission info, or discharge info, or lab result, or medication order
* Predictive: this is the most complex definition of completeness. It’s the capability to predict outcomes and care actions based on collected data. The metric for completeness is the data needed to make a prediction, and not amount of data. This means 2 different records with different data profiles could potentially be complete as long as they’re accurately predictive. Good starting point: predicting return visits

**REFERENCES**

1. *Standards for the Clinical Structure and Content of Patient Records*(Rep.). (2013, July). Retrieved June 20, 2016, from Health & Social Care Information Centre website: <http://theprsb.org/publications/bible-sets-out-the-latest-agreed-standards>
2. Statement on Quality Healthcare Data and Information. (2015). Retrieved June 20, 2016, from <http://bok.ahima.org/doc?oid=107773#.V2gGqJMrLR0>
3. Weiskopf, N. G., Hripcsak, G., Swaminathan, S., & Weng, C. (2013). Defining and measuring completeness of electronic health records for secondary use. *Journal of Biomedical Informatics,* *46*(5), 830-836. doi:10.1016/j.jbi.2013.06.010
4. Thiru, K. (2003). Systematic review of scope and quality of electronic patient record data in primary care. *Bmj,* *326*(7398), 1070-0. doi:10.1136/bmj.326.7398.1070
5. *The Six Primary Dimensions for Data Quality Assessment* (Rep.). (2013, October). Retrieved June 20, 2016, from DAMA UK Working Group website: <https://www.dqglobal.com/wp-content/uploads/2013/11/DAMA-UK-DQ-Dimensions-White-Paper-R37.pdf>
6. Integrity of the Healthcare Record: Best Practices for EHR Documentation (2013 update). (2013). Retrieved June 21, 2016, from <http://library.ahima.org/doc?oid=300257#.V2lbrpMrLAw>
7. Emergency Department Visits. (2016). Retrieved June 21, 2016, from http://www.cdc.gov/nchs/fastats/emergency-department.htm