**Meeting notes with IHPI team on Feb 13, 2019**

The corresponding slides are in DataQuestions\_zwedit\_mledit\_29jan2019.pptx

**Procedure codes**

* Claims regarding procedure codes can be submitted by either a professional or a facility.
  + The procedure codes submitted by a professional are usually CPT/HCPCS codes.
  + Facilities include labs, inpatient facilities, and outpatient facilities. Labs usually submit CPT/HCPCS codes. Inpatient facilities, such as nursing homes, and outpatient facilities, such as rehabilitation clinics, usually submit ICD procedural codes.
  + In particular, ICD procedural codes are more likely to appear in inpatient facilities than outpatient facilities.
  + Facilities follow their conventional practice to submit CPT/HCPCS codes or to submit ICD procedural codes, and there is no fixed rule.
* CPT/HCPCS codes can reflect that a lab test is ordered, but does not reflect the specific lab tests performed on a patient, which can only be seen if the lab results are submitted by labs within the United Health network. In addition, results of mundane lab tests, such as blood pressure measured in a clinic, are likely not recorded in the lab claims data.
* Ideally, a procedure performed on a patient should have two codes, one CPT/HCPCS code from a provider who orders the procedure, and one ICD procedural code from a facility which performs the action. However, this is not at all the case in the OptumInsight data. Instead, we can hardly find any ICD procedural codes in the PROC1-25 in medical claims or PROC1-5 in confinement claims.
* In order to draw a complete picture of the procedure codes, we need to pool all CPT/HCPCS and ICD procedural codes together.

**Place of service (POS)**

* The pos variable can be found in medical claims and confinement claims data.
* However, there may be errors in the pos because a service can be performed at a place different than a place the hospital may think. For example (which may be a made-up example), straightening back can only be recorded as in an emergency room (pos = 23), but in reality a doctor may straighten a patient’s back in an outpatient hospital (pos = 22).
* The IHPI team suggested we consult with medical experts to identify the errors in pos.

**Cost variables**

* The charge variable reflects the actual dollar amount a hospital or a professional asks for the insurance company.
* The reimbursement paid by the insurance is usually different than the charge, and the actual amount is masked in standard cost (std\_cost­). The standard cost is calculated through some unknown formula, and its number does not provide any information regarding hospital utility.
* The out-of-pocket amount paid by a patient is the sum of copay + coinsurance + deductible.
* To see breakdown of the confinement charges, look for rows in the medical claims with all variables matched, including patid, clmid, and conf\_id, etc. Specifically, only look at rows with Tos\_Cd = “FAC\_IP\_ACUTE”, and paid status = “P”.
* A claim in the medical table can be split into a sequence of records, in multiple rows with the same clmid. The paid status of each row indicates whether the record was declined (“D”) or paid (“P”). If all rows of the same claim were declined, then the entire claim would be removed from the medical table and the clmid would not be found.

**Confinement**

* There are some sequences of confinement claims where the lst\_dt is the same as the fst\_dt of the next claim. This may be due to transfer of location.

**Diagnoses**

* In the medical claims table, if conf\_id != “”, then we may consider the row as an inpatient record. Otherwise, we consider the row as an outpatient record, unless the fst\_dt of the row falls within the range of a confinement period.
* Using the conf\_id criterion as the previous point, there are rows with non-empty conf\_id but pos are outpatient settings. We do not need to worry about these because the pos may not reflect the correct place.
* When recording the diagnosis codes, doctors *type* in the codes. Therefore, unless a doctor is super familiar with the coding system and is an expert in all diseases, it is likely that there are coding errors such as typos and lack of specificity in the recorded codes.
* Consider the following example: 585 and 585.9 appear on the same day at place of service 21. 585 means Chronic kidney disease (ckd).

There are two possible reasons why the disease appears with two different levels of specificity.

1. The 585 is a typo.
2. The doctor was unsure about the specificity of ckd, and therefore put both trying to catch the most accurate description.

* We may ignore the diagnosis codes in the data which are not actual ICD-9 codes.
* In order to count the number of times a diagnosis code appears in a patient, we need to define an “episode of care”, depending on time, conf\_id, and patient status etc.
* Consider the following example: 1889 appears twice in the picture, in the 90st and 92rd rows. Each row has a different charge amount or place of service. If the two rows have the same clm\_id, then we may consider the two 188.9 the same.

