

Processing Claims Data in SEER*DMS

Insurance claims provide a high volume of data. Therefore, claims data processing in SEER*DMS must be automated as much as possible to avoid over-burdening registry staff. The workflow developed for claims will eventually be applied to other data streams (i.e. diagnostic indices, pathology reports).

Claims data will initially provide:

- A large amount of raw data that must be available in SEER*DMS, but not stored in the SEER*DMS RECORD table.
- Data that should be matched against itself to remove functional duplicates.
- Data that should be matched against existing patient sets and tumors to determine linkage. Links to a Patient Set and CTC are stored with the claim.
- Data that should be used to supplement treatment information for existing patient sets.

In the future, claims *may* be used:

- To identify new patients or tumors.

The processing involves:

- Loading the data from text files and storing the raw data in the PRE_RECORD table. The PRE_RECORD table consists of JSON fields. The advantage of this data structure is that a new import type could be added to the table with no structural changes to the database. IMS developed an import algorithm to parse the data and load it into JSON structures.
- Matching the data against itself to remove functional duplicates.
- Matching against and linking to existing patient data at the patient and tumor level.
- Augmenting SEER*DMS patient data with claims data. Initially, a semi-automated process will be used. A QC task will be created to review patient sets that indicate no chemo, but claims data indicate that chemo was provided. The goal is to increase the automation of this step.

Processing New Claims in SEER*DMS

