**9/20/2019 Milestone 1 Update**

Milestone 1: Understand non-compliance rate / sample sizes / PDF type mix / PDF types of each sponsor. This needs help from PHSR such as providing manually processed datasets, and summary statistics of encounter data upon request. DMG will prioritize work subject to quality of preliminary content extraction from different PDF types. Given above, DMG will provide estimate of % of automation expected (% PDF processed). By week 2.

**Data acquisition:**

* Both Chris and Luke have successfully been authorized to access to the EMTM tool on the CMS Enterprise Portal.
* We have received all supporting documentation PDFs from Q2 and Q3 2019.
* We have received all manual review result spreadsheets from Q2, Q3, and Q4 2018 + Q1, Q2 2019

**Data ingestion:**

* Successfully ingested all supporting documentation PDFs, recognizing text in the PDFs, and storing the unstructured plain text in a single CSV. Done using Python and its pytesseract library. Process takes about a minute per PDF. From some anecdotal checks, PDFs are generally legible, and word recognition accuracy is close to perfect (>99%).
* Successfully loaded and cleaned all manual review results, and stored as a single CSV. 3,899 document reviews loaded in total. Basic summary statistics of manual reviews
  + 293 of 3,899 reviews determined failure (7.5%)
  + 160 different kinds of data elements inspected; widely varying failure rates and N’s. Crosstab of failures by element included at the end in appendix, if interested.

**Data validation:**

CK2 Validation:

**Original: E-MTM Data Validation Work Plan**

**Goal**: Automate the process of validating Enhanced Medication Therapy Management (E-MTM) encounter data using contents in supporting documents provided by Medicare Part D E-MTM participants (i.e. sponsors).

**Overview of work flow:**

% encounter records

* with drug id
* identify location of service
* with valid SNOMED CT code
* etc…

Simple counting work that has been done by PHSR, no automation is needed

internal validation:  
(simple counting)

raw encounter data

primary task

find evidence of encounter in document

Scope of DMG’s automation work

external validation

detect encounter details in document content but not in encounter data

secondary task

supporting document

**Outline of Tasks:**

1. Find evidence of encounter in document
   1. Identify columns in encounter data that can be validated (ids, date of encounter, description of encounter, drug RxCUI code, etc.) using supporting document information
   2. Associate supporting document with encounters using RECORD and CNTRCT\_ID columns in encounter data, and identify rows in encounter data (encounters) that can be validated
   3. Extract contents of supporting document PDFs
      1. Extract structured textual content (e.g. tables with directly usable rows/columns - e.g. encounter date)
      2. Extract unstructured textual content (e.g. paragraphs)
      3. Extract text from images (e.g. screenshot)
      4. Validate content extraction - compare machine and human extraction results.
   4. Structuralize extracted contents as needed. This often involves identifying and preserving tables/mapping relationships between textual elements, e.g. ‘Date of Encounter’ should be associated with a date entry ‘08/26/2019’.
      1. Validate content structuralization - compare machine and human structuralization results
   5. Compare cleaned up, structuralized content data against encounter data
      1. Design a rule to identify ‘matches’ - e.g. ‘Discussed with patient’ in encounter data and ‘Talked to patient about drug use’ from supporting document content should be considered good matches. This involves the design of a semantic distance function to measure proximity of semantics expressed by different terms.
      2. Flag ‘bad matches’ - i.e. flag encounters for which our algorithm shows low likelihood of finding needed evidence in supporting documents, thus need subsequent manual curation. This requires defining a numerical threshold for the likelihood of match to flag the bad matches.
   6. Validate the flagging of mismatches
      1. Determine the minimum required sample size to train a NLP model (e.g. identifying certain encounters from a paragraph)
      2. Use a sample of manually checked encounters (with or without evidence from documents), compute performance of the matching algorithm, and determine optimal numerical threshold above.
   7. Flag sponsor compliance status (full compliance = evidence found in documents for all encounters, partial compliance = % of encounters with evidence found, etc.)
   8. Flag certain non-compliance types of interest
      1. Recommendation to change medication - does document suggest which new medication should be used?
      2. Other non-compliance types? - check with PHSR.
2. Detect encounter details in document content but not in encounter data
   1. Supervised method
      1. Check with PHSR about specific additional encounter details of interest.
      2. Identify a set of keywords, and search for presence in documents
      3. Design a scoring function to make recommendation for manual checking for encounter details
      4. Validate the recommendation algorithm - compare machine and human content extraction results
   2. Unsupervised method
      1. Identify topics embedded in text in documents
      2. In consult with PHSR, design a recommendation algorithm to flag certain topics for manual checking for encounter details
      3. Validate the recommendation algorithm - compare machine and human content extraction results.

**Appendix: Cross-tabulation of Failures by Data Element**

| **Data Element** | **Fail** | | **Pass** | | **Total** |
| --- | --- | --- | --- | --- | --- |
| **N** | **Percent** | **N** | **Percent** |
| "Drug not collected - ""too expensive"" | 6 | 85.7% | 1 | 14.3% | 7 |
| Additional medication therapy required | 2 | 2.4% | 82 | 97.6% | 84 |
| Administration of substance to produce immunity, either active or passive | 0 | 0.0% | 6 | 100.0% | 6 |
| Administration of vaccine to produce active immunity | 6 | 100.0% | 0 | 0.0% | 6 |
| Adverse drug interaction | 0 | 0.0% | 6 | 100.0% | 6 |
| Adverse drug interaction with drug | 0 | 0.0% | 6 | 100.0% | 6 |
| Adverse reaction caused by drug | 0 | 0.0% | 18 | 100.0% | 18 |
| Advice to continue with drug treatment | 0 | 0.0% | 6 | 100.0% | 6 |
| Advice to start drug treatment | 0 | 0.0% | 6 | 100.0% | 6 |
| Assessment of adverse drug reactions | 0 | 0.0% | 69 | 100.0% | 69 |
| Assessment of barriers to adherence | 0 | 0.0% | 12 | 100.0% | 12 |
| Assessment of compliance with medication regimen | 4 | 4.4% | 87 | 95.6% | 91 |
| At risk for drug therapy problem | 0 | 0.0% | 24 | 100.0% | 24 |
| At risk for negative response to medication | 0 | 0.0% | 12 | 100.0% | 12 |
| At risk for noncompliance | 3 | 9.4% | 29 | 90.6% | 32 |
| At risk of adverse drug interaction | 0 | 0.0% | 6 | 100.0% | 6 |
| At risk of medication side effect | 2 | 3.8% | 51 | 96.2% | 53 |
| Chart evaluation by healthcare professional | 2 | 7.1% | 26 | 92.9% | 28 |
| Chronic disease process education | 7 | 22.6% | 24 | 77.4% | 31 |
| Compliance issues discussed with patient | 0 | 0.0% | 21 | 100.0% | 21 |
| Comprehensive medication therapy review | 6 | 4.6% | 124 | 95.4% | 130 |
| Consultation | 6 | 50.0% | 6 | 50.0% | 12 |
| Consultation for transition of care | 1 | 2.3% | 43 | 97.7% | 44 |
| Cost effective medication alternatives available | 0 | 0.0% | 12 | 100.0% | 12 |
| Deficient knowledge of medication regimen | 0 | 0.0% | 19 | 100.0% | 19 |
| Diabetic care education | 0 | 0.0% | 12 | 100.0% | 12 |
| Diabetic education | 2 | 10.5% | 17 | 89.5% | 19 |
| Difficulty following instructions | 0 | 0.0% | 12 | 100.0% | 12 |
| Discharged from hospital | 0 | 0.0% | 6 | 100.0% | 6 |
| Discussed with doctor | 1 | 2.6% | 37 | 97.4% | 38 |
| Discussed with patient | 0 | 0.0% | 74 | 100.0% | 74 |
| Discussed with prescriber | 1 | 5.3% | 18 | 94.7% | 19 |
| Discussion about refilling prescription | 0 | 0.0% | 28 | 100.0% | 28 |
| Documentation of current medications | 0 | 0.0% | 18 | 100.0% | 18 |
| Documentation of medication related action plan | 2 | 6.1% | 31 | 93.9% | 33 |
| Does not refill medications appropriately | 0 | 0.0% | 28 | 100.0% | 28 |
| Drug compliance good | 0 | 0.0% | 7 | 100.0% | 7 |
| Drug dose | 0 | 0.0% | 6 | 100.0% | 6 |
| Drug interaction | 0 | 0.0% | 24 | 100.0% | 24 |
| Drug intolerance | 0 | 0.0% | 6 | 100.0% | 6 |
| Drug not recommended in elderly | 0 | 0.0% | 28 | 100.0% | 28 |
| Drug therapy discontinued | 1 | 2.3% | 42 | 97.7% | 43 |
| Duration of medication therapy too long | 0 | 0.0% | 12 | 100.0% | 12 |
| Dyslipidemia education | 0 | 0.0% | 6 | 100.0% | 6 |
| Education | 1 | 8.3% | 11 | 91.7% | 12 |
| Education provision | 0 | 0.0% | 28 | 100.0% | 28 |
| Evaluation procedure | 4 | 16.7% | 20 | 83.3% | 24 |
| Failed encounter | 0 | 0.0% | 28 | 100.0% | 28 |
| Follow-up consultation | 0 | 0.0% | 6 | 100.0% | 6 |
| Health literacy assessment | 6 | 50.0% | 6 | 50.0% | 12 |
| Healthcare information exchange | 0 | 0.0% | 6 | 100.0% | 6 |
| High risk of adverse drug event | 9 | 33.3% | 18 | 66.7% | 27 |
| Hypertension education | 0 | 0.0% | 6 | 100.0% | 6 |
| Hypertensive disorder, systemic arterial | 3 | 42.9% | 4 | 57.1% | 7 |
| Immunization advised | 0 | 0.0% | 13 | 100.0% | 13 |
| Inappropriate medication stopped | 0 | 0.0% | 6 | 100.0% | 6 |
| Inbound communication | 3 | 50.0% | 3 | 50.0% | 6 |
| Influenza immunization advised | 0 | 0.0% | 7 | 100.0% | 7 |
| Information gathering | 0 | 0.0% | 21 | 100.0% | 21 |
| Informing doctor | 4 | 33.3% | 8 | 66.7% | 12 |
| Initial patient assessment | 3 | 15.8% | 16 | 84.2% | 19 |
| Letter Encounter to Patient | 0 | 0.0% | 25 | 100.0% | 25 |
| Medication Reconciliation | 0 | 0.0% | 6 | 100.0% | 6 |
| Medication administration education | 0 | 0.0% | 13 | 100.0% | 13 |
| Medication commenced | 6 | 50.0% | 6 | 50.0% | 12 |
| Medication dose too high | 0 | 0.0% | 18 | 100.0% | 18 |
| Medication dosing interval changed | 7 | 100.0% | 0 | 0.0% | 7 |
| Medication education | 0 | 0.0% | 59 | 100.0% | 59 |
| Medication interaction education | 0 | 0.0% | 12 | 100.0% | 12 |
| Medication interaction with medication education | 6 | 19.4% | 25 | 80.6% | 31 |
| Medication prescription education, guidance and counseling | 3 | 8.6% | 32 | 91.4% | 35 |
| Medication reconciliation | 0 | 0.0% | 33 | 100.0% | 33 |
| Medication regimen compliance education | 3 | 4.8% | 59 | 95.2% | 62 |
| Medication review declined | 4 | 16.0% | 21 | 84.0% | 25 |
| Medication side effects education | 5 | 6.7% | 70 | 93.3% | 75 |
| Medication side effects present | 0 | 0.0% | 6 | 100.0% | 6 |
| Medication therapy changed | 0 | 0.0% | 6 | 100.0% | 6 |
| Medication therapy management information sent | 0 | 0.0% | 56 | 100.0% | 56 |
| Medication therapy management information sent to health care provider | 12 | 14.6% | 70 | 85.4% | 82 |
| Medication therapy management recommendation accepted by patient | 1 | 3.6% | 27 | 96.4% | 28 |
| Medication therapy management recommendation accepted by prescriber | 5 | 16.1% | 26 | 83.9% | 31 |
| Medication therapy management recommendation acknowledged by patient | 0 | 0.0% | 6 | 100.0% | 6 |
| Medication therapy management recommendation acknowledged by prescriber | 0 | 0.0% | 6 | 100.0% | 6 |
| Medication therapy management recommendation refused by patient | 5 | 7.5% | 62 | 92.5% | 67 |
| Medication therapy management recommendation refused by prescriber | 2 | 5.6% | 34 | 94.4% | 36 |
| Medication therapy unnecessary | 0 | 0.0% | 6 | 100.0% | 6 |
| Met sponsor's approved criteria for $0 cost sharing assistance | 0 | 0.0% | 14 | 100.0% | 14 |
| Met sponsor's auto-referral targeting criteria | 0 | 0.0% | 20 | 100.0% | 20 |
| More effective medication therapy available | 8 | 22.2% | 28 | 77.8% | 36 |
| Multiple chronic diseases | 0 | 0.0% | 13 | 100.0% | 13 |
| Multiple medications taken for condition appropriately treated with single medication therapy | 2 | 5.1% | 37 | 94.9% | 39 |
| New patient questionnaire completed | 1 | 8.3% | 11 | 91.7% | 12 |
| No drug therapy prescribed | 0 | 0.0% | 6 | 100.0% | 6 |
| No known drug allergy | 0 | 0.0% | 6 | 100.0% | 6 |
| No response to Medication therapy management recommendation by prescriber | 0 | 0.0% | 6 | 100.0% | 6 |
| Non-compliance of drug therapy | 2 | 14.3% | 12 | 85.7% | 14 |
| Noncompliance with medication regimen | 2 | 4.3% | 44 | 95.7% | 46 |
| Not taking medication for chronic disease | 1 | 3.6% | 27 | 96.4% | 28 |
| Over-the-counter medication started | 0 | 0.0% | 6 | 100.0% | 6 |
| Patient did not respond to medication therapy management outreach attempts | 0 | 0.0% | 25 | 100.0% | 25 |
| Patient enrolled in program | 0 | 0.0% | 7 | 100.0% | 7 |
| Patient forgets to take medication | 0 | 0.0% | 6 | 100.0% | 6 |
| Patient medication education | 3 | 15.8% | 16 | 84.2% | 19 |
| Patient notified | 12 | 34.3% | 23 | 65.7% | 35 |
| Patient notified of eligibility for medication therapy management service | 3 | 3.9% | 74 | 96.1% | 77 |
| Patient opts out of enhanced medication therapy management services | 0 | 0.0% | 6 | 100.0% | 6 |
| Patient registration | 0 | 0.0% | 12 | 100.0% | 12 |
| Potential drug-disease interaction | 1 | 3.6% | 27 | 96.4% | 28 |
| Potential drug-drug interaction | 0 | 0.0% | 24 | 100.0% | 24 |
| Prescription medication discontinued | 0 | 0.0% | 7 | 100.0% | 7 |
| Prescription medication started | 0 | 0.0% | 36 | 100.0% | 36 |
| Preventive medication therapy needed | 0 | 0.0% | 39 | 100.0% | 39 |
| Promotion of adherence to medication | 0 | 0.0% | 18 | 100.0% | 18 |
| Provided medication at $2 Member incentive | 0 | 0.0% | 14 | 100.0% | 14 |
| Provision of medication list | 1 | 7.7% | 12 | 92.3% | 13 |
| Provision of medication related action plan | 4 | 4.2% | 91 | 95.8% | 95 |
| Recommendation accepted | 0 | 0.0% | 20 | 100.0% | 20 |
| Recommendation acknowledged by patient | 0 | 0.0% | 6 | 100.0% | 6 |
| Recommendation to change medication | 5 | 6.1% | 77 | 93.9% | 82 |
| Recommendation to change medication dosing interval | 6 | 100.0% | 0 | 0.0% | 6 |
| Recommendation to change medication timing of administration | 0 | 0.0% | 6 | 100.0% | 6 |
| Recommendation to continue a medication | 0 | 0.0% | 6 | 100.0% | 6 |
| Recommendation to continue with drug treatment | 1 | 2.9% | 34 | 97.1% | 35 |
| Recommendation to decrease medication dose | 0 | 0.0% | 12 | 100.0% | 12 |
| Recommendation to discontinue medication | 1 | 4.2% | 23 | 95.8% | 24 |
| Recommendation to discontinue over-the-counter medication | 0 | 0.0% | 6 | 100.0% | 6 |
| Recommendation to discontinue prescription medication | 1 | 4.2% | 23 | 95.8% | 24 |
| Recommendation to monitor medication adherence | 0 | 0.0% | 6 | 100.0% | 6 |
| Recommendation to monitor physiologic parameters | 0 | 0.0% | 7 | 100.0% | 7 |
| Recommendation to start drug treatment | 1 | 4.8% | 20 | 95.2% | 21 |
| Recommendation to start medication monitoring | 0 | 0.0% | 47 | 100.0% | 47 |
| Recommendation to start medication therapy | 0 | 0.0% | 27 | 100.0% | 27 |
| Recommendation to start over-the-counter medication | 1 | 3.3% | 29 | 96.7% | 30 |
| Recommendation to start prescription medication | 2 | 2.2% | 91 | 97.8% | 93 |
| Recommendation to stop drug treatment | 2 | 6.1% | 31 | 93.9% | 33 |
| Referral to pharmacist | 0 | 0.0% | 6 | 100.0% | 6 |
| Referral to physician | 12 | 63.2% | 7 | 36.8% | 19 |
| Referred by payer | 4 | 28.6% | 10 | 71.4% | 14 |
| Referred for medication therapy management | 3 | 14.3% | 18 | 85.7% | 21 |
| Risk Assessment | 5 | 17.9% | 23 | 82.1% | 28 |
| Risk evaluation and mitigation strategy consultation | 0 | 0.0% | 18 | 100.0% | 18 |
| Suspected non-compliance of drug therapy | 1 | 1.3% | 75 | 98.7% | 76 |
| Synchronization of repeat medication | 0 | 0.0% | 6 | 100.0% | 6 |
| Taking high risk medication | 0 | 0.0% | 6 | 100.0% | 6 |
| Taking medication for chronic disease | 13 | 52.0% | 12 | 48.0% | 25 |
| Taking multiple medications for chronic disease | 10 | 23.8% | 32 | 76.2% | 42 |
| Targeted medication therapy review | 9 | 8.6% | 96 | 91.4% | 105 |
| Telemedicine consultation with patient | 5 | 27.8% | 13 | 72.2% | 18 |
| Telephone Encounter | 1 | 3.6% | 27 | 96.4% | 28 |
| Telephone consultation | 8 | 44.4% | 10 | 55.6% | 18 |
| Telephone encounter | 1 | 16.7% | 5 | 83.3% | 6 |
| Telephone follow-up | 2 | 10.5% | 17 | 89.5% | 19 |
| Therapy changed to alternate medication | 5 | 83.3% | 1 | 16.7% | 6 |
| Transition of Care | 0 | 0.0% | 7 | 100.0% | 7 |
| Transition of care | 6 | 33.3% | 12 | 66.7% | 18 |
| Unable to reach | 3 | 10.7% | 25 | 89.3% | 28 |
| Under care of multiple providers | 0 | 0.0% | 13 | 100.0% | 13 |
| Unmet drug need identified | 0 | 0.0% | 12 | 100.0% | 12 |
| Uses multiple pharmacies | 1 | 14.3% | 6 | 85.7% | 7 |
| Verification of allergy status | 0 | 0.0% | 12 | 100.0% | 12 |