Appendix A - Example of the Prompt for the TASK-1

Part A of the final Prompt - Task definition

You will receive text that may contain one or more data tables.

It describes a Cohort for OHDSI OMOP.

Your task:

1. Identify all tables in the input that reflect Concept Sets.

2. Extract each table separately.

3. Return them in strict JSON with this structure:

```
{
 "tables": [
   {
    "table_id": "table_1",
    "name": "Try to extract table name from the title…",
    "rows": [
      {
          "Concept_ID": "...",
          "Concept_Name": "…",
          "Code": "…",
          "Vocabulary": "…",
          "Excluded": "…",
          "Descendants": "…",
          "Mapped": "…"
      },
      {
          …
      }]},
  {
    "table_id": "table_2",
    "name": ""Try to extract table name from the title…"",
    "rows": [
      {...}
    ]}]}
```

Part B of the final Prompt - Restrictions

Do not merge tables.

Do not change column names.

Do not create extra explanations.

Only valid JSON.

STRICT MODE:

Return only the final output in the format below. No reasoning, no commentary.

Part C of the final Prompt - Added parameters

Data

<<<

 # Cohort Definition

Human Readable Cohort Definition

Cohort Entry Events

People with continuous observation of 365 days before event may enter the cohort when observing any of the following:

1. drug exposure of '[HowOften] IL23 inhibitors' for the first time in the person's history.

Limit cohort entry events to the earliest event per person.

Inclusion Criteria

1. nested in Plaque psoriasis

Entry events having at least 1 condition occurrence of 'Plaque psoriasis', starting anytime on or before cohort entry start date.

Cohort Exit

The cohort end date will be based on a continuous exposure to '[HowOften] IL23 inhibitors':

allowing 90 days between exposures, adding 0 days after exposure ends, and using days supply and exposure end date for exposure duration.

Cohort Eras

Entry events will be combined into cohort eras if they are within 0 days of each other.

Concept sets:

[HowOften] IL23 inhibitors

|Concept ID|Concept Name|Code|Vocabulary|Excluded|Descendants|Mapped

|:---|:----------|:--|:-----|:--:|:--:|:--:|

|35200139|tildrakizumab|2053436|RxNorm|NO|YES|NO|

|1511348|risankizumab|2166040|RxNorm|NO|YES|NO|

|1593700|guselkumab|1928588|RxNorm|NO|YES|NO|

Plaque psoriasis

|Concept ID|Concept Name|Code|Vocabulary|Excluded|Descendants|Mapped

|:---|:----------|:--|:-----|:--:|:--:|:--:|

|140168|Psoriasis|9014002|SNOMED|NO|NO|NO|

|4307925|Psoriasis vulgaris|200975007|SNOMED|NO|YES|NO|

|4063431|Plaque psoriasis|200965009|SNOMED|NO|YES|NO|

>>>

Appendix B - Example of the Prompt for the TASK-2

Part A of the final Prompt - Task definition

You will receive a text that may describe a Cohort for OHDSI OMOP.

Your task:

1. Extract information from the Cohort description

2. Process each <<<TODO: >>> as a separate instruction and replace it with the asked value

3. Check Cohort Entry Events in Cohort Description:

- Add ConditionOccurrence block only if it is existing there

- Add Observation block only if it is existing there

- Add DrugExposure block only if it is existing there

4. Return a JSON with this structure:

Part B of the final Prompt - Desired JSON schema with inline instructions

```
{
  "cdmVersionRange" : ">=5.0.0",
  "PrimaryCriteria" : {
    "CriteriaList" : [
      {
        "ConditionOccurrence" : {
          "CodesetId" : <<<TODO:
```

1. Search in the Cohort Description for the name of the Condition and then identify its "id" from "Concept Set Data JSON".

2. Insert that id here

3. Do not use CONCEPT_ID. Use only "id"

>>>,

"ConditionTypeExclude" : false

}},

{

"Observation" : {

"CodesetId" : <<<TODO:

1. Search in the Cohort Description for the name of Observation and then identify its id from "Concept Set Data JSON".

2. Insert that id here

3. Do not use CONCEPT_ID. Use only "id"

>>>,

"ObservationTypeExclude" : false

}},

{

"DrugExposure" : {

"CodesetId" : <<<TODO:

1. Search in the Cohort Description for the name of Drug Exposure and then identify its id from "Concept Set Data JSON".

2. Insert that id here

3. Do not use CONCEPT_ID. Use only "id"

>>>,

"First" : <<<TODO: Identify True or False>>>,

"DrugTypeExclude" : false

}}],

"ObservationWindow" : {

"PriorDays" : <<<TODO:

1. in the Cohort Description, identify the prior days for observation

2. Use only information from Cohort Entry Events. Ignore the rest of the information

3. If you can't find any, put 0

>>>,

"PostDays" : <<<TODO: in Cohort Description, identify the post days for observation

1. in the Cohort Description, identify the post days for observation

2. Use only information from Cohort Entry Events. Ignore the rest of the information

3. If you can't find any, put 0

>>>

},

"PrimaryCriteriaLimit" : {

"Type" : <<<TBD: Identify, "First" or "All". If you can't find a mention of First, then All>>>

}}}

Part C of the final Prompt - Restrictions

Do not change column names.

Do not create extra explanations.

Only valid JSON.

STRICT MODE:

Return only the final output in the format below. No reasoning, no commentary.

Part D of the final Prompt - Added parameters (Cohort Definition)

Cohort Description:

<<<

 # Cohort Definition

Human Readable Cohort Definition

Cohort Entry Events

People enter the cohort when observing any of the following:

1. condition occurrences of 'Diarrhea'.

Cohort Exit

The cohort end date will be offset from index event's end date plus 7 days.

Cohort Eras

Entry events will be combined into cohort eras if they are within 0 days of each other.

Concept sets:

Diarrhea

|Concept ID|Concept Name|Code|Vocabulary|Excluded|Descendants|Mapped

|:---|:----------|:--|:-----|:---:|:--:|:--:|

|196523|Diarrhea|62315008|SNOMED|NO|YES|NO|

|80141|Functional diarrhea|47812002|SNOMED|NO|YES|NO|

>>>

Part E of the final Prompt - Added parameters (Concept Sets from the TASK-1)

Concept Set Data JSON:

```
<<<
[
    {
      "id" : 1,
      "name" : "Diarrhea",
      "expression" : {
        "items" : [
          {
            "concept" : {
              "CONCEPT_ID" : 196523,
              "CONCEPT_NAME" : "Diarrhea",
              "STANDARD_CONCEPT" : "S",
              "STANDARD_CONCEPT_CAPTION" : "Standard",
              "INVALID_REASON" : "V",
              "INVALID_REASON_CAPTION" : "Valid",
              "CONCEPT_CODE" : "62315008",
              "DOMAIN_ID" : "Condition",
              "VOCABULARY_ID" : "SNOMED",
              "CONCEPT_CLASS_ID" : "Clinical Finding"
            },
            "isExcluded" : false,
            "includeDescendants" : true,
            "includeMapped" : false
          },
          {
            "concept" : {
              "CONCEPT_ID" : 80141,
              "CONCEPT_NAME" : "Functional diarrhea",
              "STANDARD_CONCEPT" : "S",
              "STANDARD_CONCEPT_CAPTION" : "Standard",
              "INVALID_REASON" : "V",
              "INVALID_REASON_CAPTION" : "Valid",
              "CONCEPT_CODE" : "47812002",
              "DOMAIN_ID" : "Condition",
              "VOCABULARY_ID" : "SNOMED",
```

```
            "CONCEPT_CLASS_ID" : "Clinical Finding"
          },
          "isExcluded" : false,
          "includeDescendants" : true,
          "includeMapped" : false
        }]}}]
>>>
```