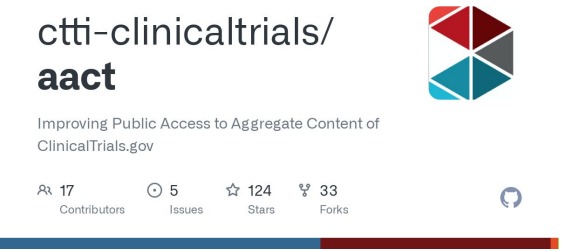
# Standardization of Clinical Trials Data for Large-scale Incidence Calculations of Adverse Drug Events

OHDSI Community Call

July 16, 2024

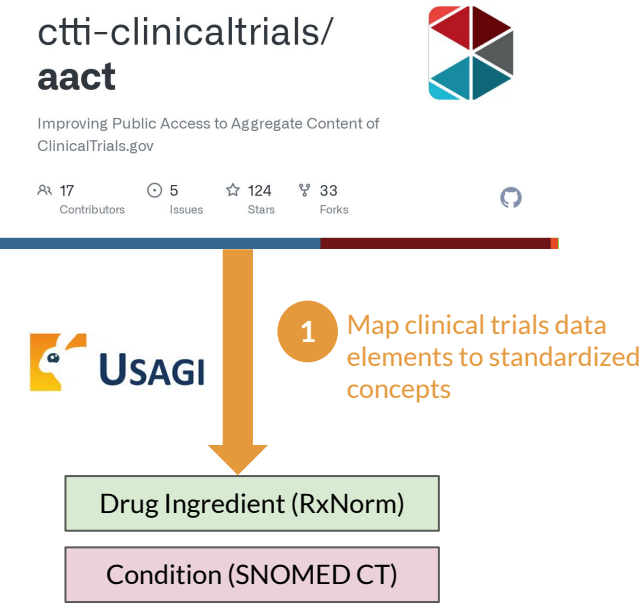
**Elise Ruan**, MD, MPH

# Clinical Trials Adverse Event Drug Data

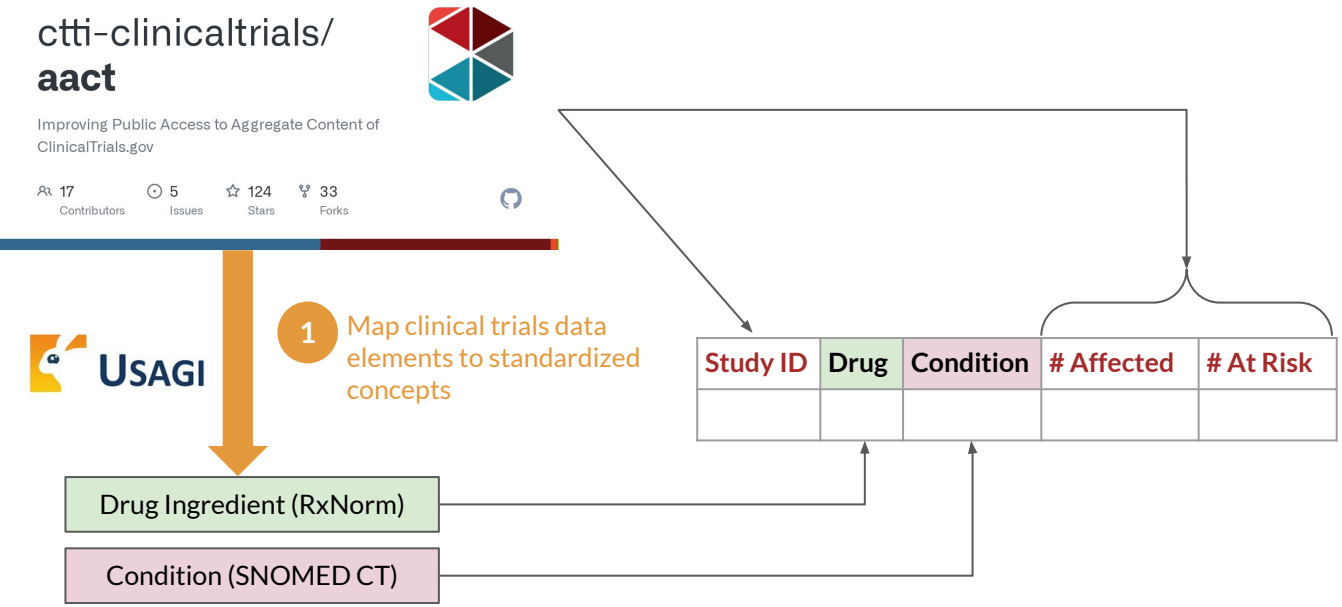



- Much of knowledge regarding potential adverse drug events (ADE) comes from premarketing clinical trials
- **Adverse event** is defined as any **unfavorable or unintended** sign, symptom, or disease **temporally** associated with the use of a drug, **without any judgement about causality** or relationship to the drug.
- Reported events (as well as most other data elements) are entered as unstructured text entries in the clinical trials database

# Objectives

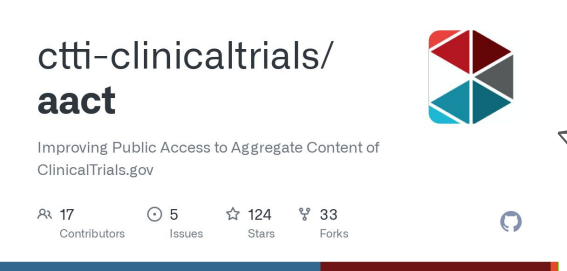

ctti-clinicaltrials/
aact
Improving Public Access to Aggregate Content of ClinicalTrials.gov
17 Contributors
5 Issues
124 Stars
33 Forks
USAGI
1
Map clinical trials data elements to standardized concepts
Drug Ingredient (RxNorm)
Condition (SNOMED CT)

# Objectives




2 Large-scale calculation of incidences for drug-condition pairs

# Objectives





1 Map clinical trials data elements to standardized concepts

Drug Ingredient (RxNorm)

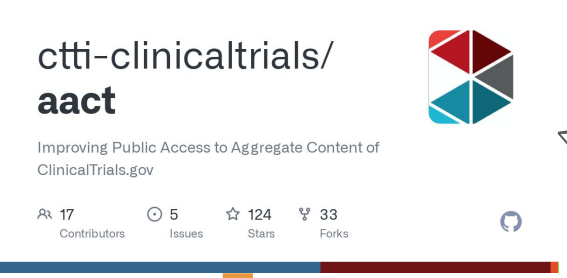
Condition (SNOMED CT)

Time at Risk (days)

| Study ID | Drug | Condition | # Affected | # At Risk | TAR |
|---|---|---|---|---|---|
| | | | | | |

2 Large-scale calculation of incidences for drug-condition pairs

# Objectives





1 Map clinical trials data elements to standardized concepts

Drug Ingredient (RxNorm)

Condition (SNOMED CT)

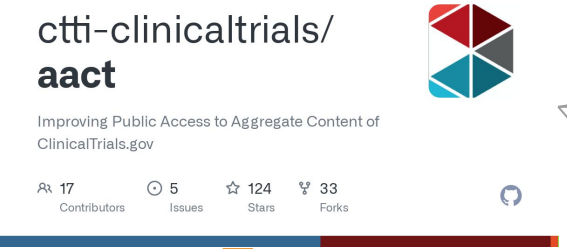
Time at Risk (days)

| Study ID | Drug | Condition | # Affected | # At Risk | TAR |
|---|---|---|---|---|---|
| | | | | | |

2 Large-scale calculation of incidences for drug-condition pairs

| Drug | Condition | Pooled Incidence | Prediction Interval |
|---|---|---|---|
| | | | |

# Objectives





1 Map clinical trials data elements to standardized concepts

Drug Ingredient (RxNorm)

Condition (SNOMED CT)

Time at Risk (days)

| Study ID | Drug | Condition | # Affected | # At Risk | TAR |
|---|---|---|---|---|---|
| | | | | | |

2 Large-scale calculation of incidences for drug-condition pairs

| Drug | Condition | Pooled Incidence | Prediction Interval | Total Variance | Interstudy Heterogeneity |
|---|---|---|---|---|---|
| | | | | | |

3 Describe variance and interstudy heterogeneity of calculated incidences

# Mapping of Drugs and Adverse Events

**457,254** Trials 2000-6/2023 → **7,543** Trials → **4,098** Trials

- Had reported adverse event data
- Intervention type was drug
- FDA approved drug

- Had only 1 intervention
- Excluded result groups that contained “placebo”

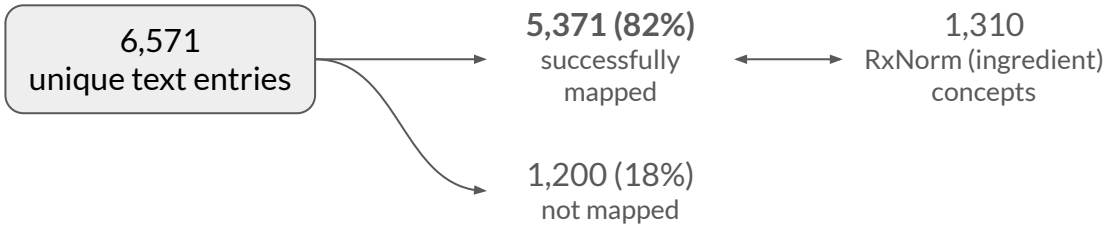
## What is AACT?

AACT is a publicly available relational database that contains all information (protocol and result data elements) about every study registered in ClinicalTrials.gov. Content is downloaded from ClinicalTrials.gov daily and loaded into AACT. The Clinical Trials Transformation Initiative (CTTI) enhanced AACT in October, 2016 to include the following features:

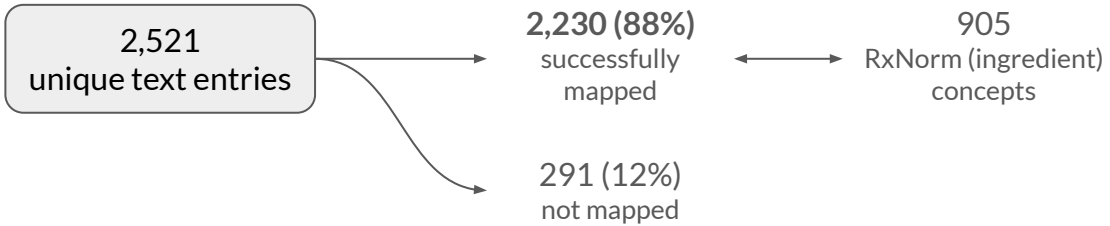
- Database content refreshed daily
- Database directly accessible in the cloud
- Static copies of the database available for download
- Open source tools freely available (postgreSQL, Ruby on Rails, Tableau Public)
- Source code available via **Github**

# Mapping Results: Drug Concepts

All Trials (n= 7,543)




Trials with 1 Intervention (n=4,098)

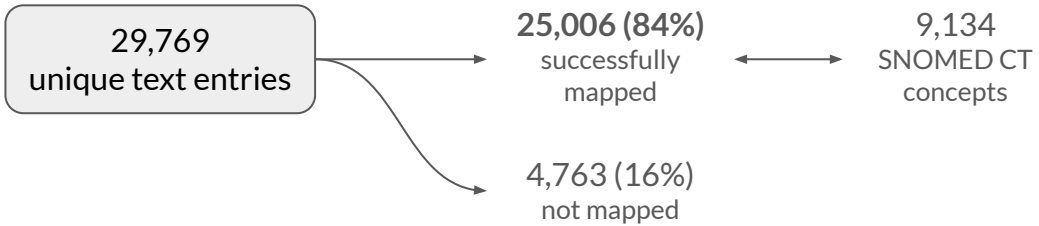



**Drug Ingredients with Most Trials**

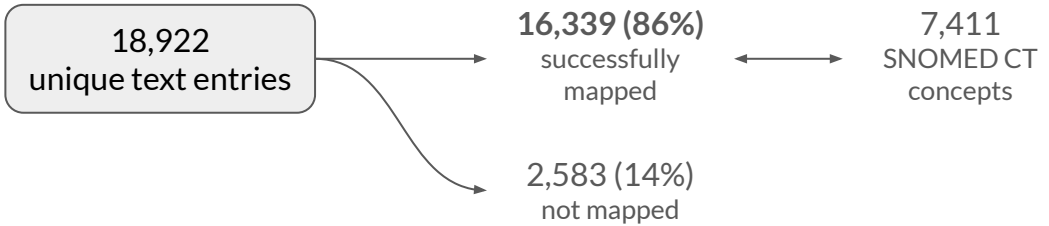
1. Cyclophosphamide
2. Pembrolizumab
3. Carboplatin
4. Dexamethasone
5. Paclitaxel
6. Nivolumab
7. Fludarabine
8. Cisplatin
9. Gemcitabine
10. Sodium Chloride

# Mapping Results: Condition Concepts

All Trials (n= 7,543)



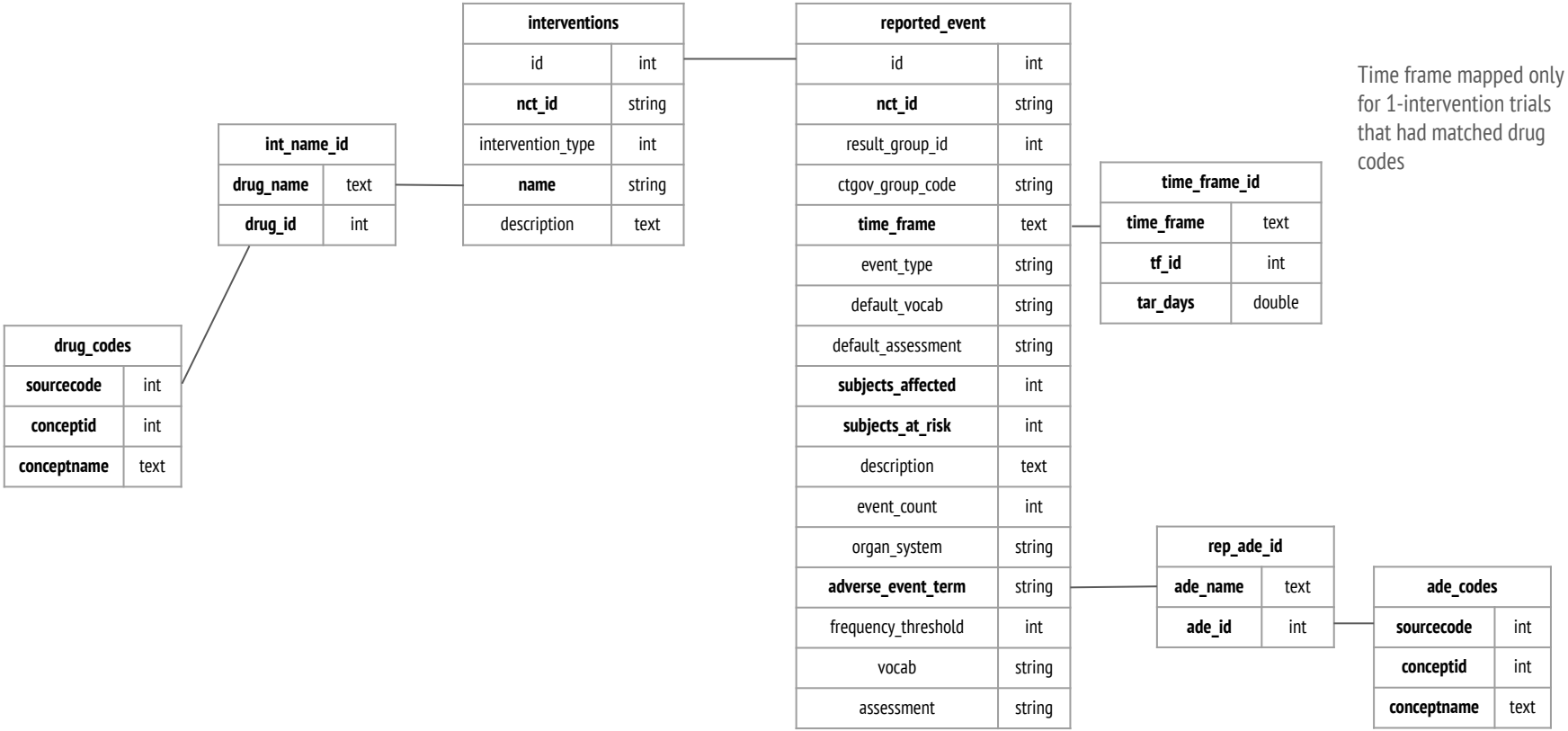

Trials with 1 Intervention (n=4,098)




**Conditions with Most Trials**

1. Headache
2. Nausea
3. Diarrhea
4. Vomiting
5. Fatigue
6. Dizziness
7. Constipation
8. Fever
9. Abdominal Pain
10. Cough

# Data Schema




Time frame mapped only for 1-intervention trials that had matched drug codes

# Incidences & Prediction Intervals

| Drug | Condition | Trial | Reported | n |
|---|---|---|---|---|
| Tramadol | Headache | 1 | 22 | 273 |
| Tramadol | Nausea | 1 | 57 | 273 |
| Tramadol | Nausea | 2 | 72 | 251 |
| Dronabinol | Nausea | 3 | 2 | 35 |
| Dronabinol | Nausea | 4 | 8 | 16 |
| Prednisone | Heartburn | 5 | 5 | 67 |
| Prednisone | Heartburn | 6 | 7 | 68 |

# Incidences & Prediction Intervals

**Incidence (Y):** Reported/ n

**Variance (V):** Y*(1-Y)/n

| Drug | Condition | Trial | Reported | n | Y | V |
|---|---|---|---|---|---|---|
| Tramadol | Headache | 1 | 22 | 273 | 0.08 | 0.0003 |
| Tramadol | Nausea | 1 | 57 | 273 | 0.21 | 0.0006 |
| Tramadol | Nausea | 2 | 72 | 251 | 0.29 | 0.0008 |
| Dronabinol | Nausea | 3 | 2 | 35 | 0.06 | 0.0015 |
| Dronabinol | Nausea | 4 | 8 | 16 | 0.50 | 0.0156 |
| Prednisone | Heartburn | 5 | 5 | 67 | 0.07 | 0.0010 |
| Prednisone | Heartburn | 6 | 7 | 68 | 0.10 | 0.0013 |

# Incidences & Prediction Intervals

**Incidence (Y):** Reported/ n

**Variance (V):** Y*(1-Y)/n

| Drug | Condition | Trial | Reported | n | Y | V |
|---|---|---|---|---|---|---|
| Tramadol | Headache | 1 | 22 | 273 | 0.08 | 0.0003 |
| Tramadol | Nausea | 1 | 57 | 273 | 0.21 | 0.0006 |
| Tramadol | Nausea | 2 | 72 | 251 | 0.29 | 0.0008 |
| Dronabinol | Nausea | 3 | 2 | 35 | 0.06 | 0.0015 |
| Dronabinol | Nausea | 4 | 8 | 16 | 0.50 | 0.0156 |
| Prednisone | Heartburn | 5 | 5 | 67 | 0.07 | 0.0010 |
| Prednisone | Heartburn | 6 | 7 | 68 | 0.10 | 0.0013 |

For pairs with only 1 study, just report incidence

RMA* (Yi, Vi)

RMA* (Yi, Vi)

RMA* (Yi, Vi)

| Drug | Condition | Incidence (Pred Interval) |
|---|---|---|
| Tramadol | Headache | 0.08 |
| Tramadol | Nausea | 0.25 (0.14, 0.39) |
| Dronabinol | Nausea | 0.21 (0.003, 0.96) |
| Prednisone | Heartburn | 0.09 (0.05, 0.15) |

*For RMA used logit(Yi) for input and then inv.logit of pooled incidence and prediction intervals

# Incidences & Prediction Intervals

| drug_conceptid | drug_conceptname | ade_conceptid | ade_conceptname | pooled_incidence | predict_lower | predict_upper |
|---|---|---|---|---|---|---|
| 701322 | memantine | 31967 | Nausea | .12827 | .0333714 | .3854285 |
| 701322 | memantine | 75860 | Constipation | .1641708 | .0584904 | .3830999 |
| 701322 | memantine | 81902 | Urinary tract infectious disease | .0201333 | .0108633 | .0370178 |
| 701322 | memantine | 132797 | Sepsis | .0276103 | .0164179 | .0460755 |
| 701322 | memantine | 133280 | Alopecia | .0886829 | .0123903 | .4301427 |
| 701322 | memantine | 134736 | Backache | .0911835 | .0296742 | .2476498 |
| 701322 | memantine | 135360 | Syncope | .0065743 | .0002753 | .137233 |
| 701322 | memantine | 138525 | Pain in limb | .0469607 | .0081583 | .2279072 |
| 701322 | memantine | 140214 | Eruption | .0314289 | .0091214 | .1026403 |
| 701322 | memantine | 196523 | Diarrhea | .0911387 | .0189835 | .3419538 |
| 701322 | memantine | 197672 | Urinary incontinence | .006162 | .0003358 | .1026838 |

# Variance and Heterogeneity of Incidences

RMA (logit(Yi), logit(Vi)) ⟶ on logit scale

| Drug | Condition | Incidence (Pred Interval) | tau2* | $v_T$** | Total Variance |
|---|---|---|---|---|---|
| Tramadol | Nausea | -1.12 (-1.78, -0.46) | 0.07 | 0.02 | 0.09 |
| Dronabinol | Nausea | -1.35 (-5.95, 3.25) | 3.54 | 0.39 | 3.93 |
| Prednisone | Heartburn | -2.31 (-2.91, -1.71) | 0 | 0.19 | 0.19 |

inv.logit()

Linear scale

| Drug | Condition | Incidence (Pred Interval) |
|---|---|---|
| Tramadol | Nausea | 0.25 (0.14, 0.39) |
| Dronabinol | Nausea | 0.21 (0.003, 0.96) |
| Prednisone | Heartburn | 0.09 (0.05, 0.15) |

*tau2 = interstudy heterogeneity
**$v_T$= typical sampling variance
Total Variance = tau2 + vt ⟶ used for prediction interval

# Variance and Heterogeneity - by Condition

RMA (logit(Yi), logit(Vi)) ⟶ on logit scale

| Drug | Condition | tau2* | $v_T$** | Total Variance |
|---|---|---|---|---|
| Vincristine | Heartburn | 0 | 0.19 | 0.19 |
| Cyclophosphamide | Heartburn | 0 | 0.19 | 0.19 |
| Rituximab | Heartburn | 0 | 0.30 | 0.30 |
| Doxorubicin | Heartburn | 0 | 0.19 | 0.19 |
| Metformin | Heartburn | 0.87 | 0.93 | 1.80 |
| Prednisone | Heartburn | 0 | 0.19 | 0.19 |
| Lenalidomide | Heartburn | 2.57 | 0.65 | 3.22 |

| Condition | Median(tau2) | Median(TV) |
|---|---|---|
| Heartburn | 0 | 0.19 |
| Hemarthrosis | 2.41 | 6.90 |

# Variance and Heterogeneity - Ex. Hemarthrosis

| Drug | Condition | tau2* | $v_T$** | Total Variance |
|---|---|---|---|---|
| Vitamin D3 | Hemarthrosis | 8.94 | 0.50 | 9.44 |
| Empagliflozin | Hemarthrosis | 3.36 | 1.00 | 4.37 |

| Condition | Median(tau2) | Median(TV) |
|---|---|---|
| Hemarthrosis | 2.41 | 6.90 |

# Variance and Heterogeneity - Ex. Hemarthrosis

| Drug | Condition | Trial | Reported | n |
|---|---|---|---|---|
| Vitamin D3 | Hemarthrosis | 1 | 2 | 22,374 |
| Vitamin D3 | Hemarthrosis | 2 | 2 | 292 |
| Empagliflozin | Hemarthrosis | 3 | 1 | 157 |
| Empagliflozin | Hemarthrosis | 4 | 1 | 2,996 |

| Drug | Condition | tau2* | $v_T$** | Total Variance |
|---|---|---|---|---|
| Vitamin D3 | Hemarthrosis | 8.94 | 0.50 | 9.44 |
| Empagliflozin | Hemarthrosis | 3.36 | 1.00 | 4.37 |

| Condition | Median(tau2) | Median(TV) |
|---|---|---|
| Hemarthrosis | 2.41 | 6.90 |

# Variance and Heterogeneity - by Drug

RMA (logit(Yi), logit(Vi)) ⟶ on logit scale

| Drug | Condition | tau2* | $v_T$** | Total Variance |
|---|---|---|---|---|
| Tramadol | Nausea | 0.07 | 0.02 | 0.09 |
| Tramadol | Constipation | 0 | 0.09 | 0.09 |
| Tramadol | Vomiting | 0.15 | 0.03 | 0.18 |
| Tramadol | Hypoxia | 0.79 | 0.13 | 0.92 |
| Tramadol | Injection Site Pain | 0 | 0.17 | 0.07 |

| Drug | Median(tau2) | Median(TV) |
|---|---|---|
| Tramadol | 0.07 | 0.09 |
| Dronabinol | 8.77 | 9.53 |

# Variance and Heterogeneity - by Drug

RMA (logit(Yi), logit(Vi)) ⟶ on logit scale

| Drug | Condition | tau2* | $v_T$** | Total Variance |
|---|---|---|---|---|
| Tramadol | Nausea | 0.07 | 0.02 | 0.09 |
| Tramadol | Constipation | 0 | 0.09 | 0.09 |
| Tramadol | Vomiting | 0.15 | 0.03 | 0.18 |
| Tramadol | Hypoxia | 0.79 | 0.13 | 0.92 |
| Tramadol | Injection Site Pain | 0 | 0.17 | 0.07 |

| Drug | Median(tau2) | Median(TV) |
|---|---|---|
| Tramadol | 0.07 | 0.09 |
| Dronabinol | 8.77 | 9.53 |

| Drug | Condition | Trial | Reported | n |
|---|---|---|---|---|
| Dronabinol | Nausea | 3 | 2 | 35 |
| Dronabinol | Nausea | 4 | 8 | 16 |

# Challenges & Limitations

- Low proportion of included studies
- Low number of studies per pair
- Combination drugs - for heartburn - 2 trials for R-CHOP - artificially brings down median

| Drug | Condition | tau2* | $v_T$** | Total Variance |
|---|---|---|---|---|
| Vincristine | Heartburn | 0 | 0.19 | 0.19 |
| Cyclophosphamide | Heartburn | 0 | 0.19 | 0.19 |
| Rituximab | Heartburn | 0 | 0.30 | 0.30 |
| Doxorubicin | Heartburn | 0 | 0.19 | 0.19 |
| Metformin | Heartburn | 0.87 | 0.93 | 1.80 |
| Prednisone | Heartburn | 0 | 0.19 | 0.19 |
| Lenalidomide | Heartburn | 2.57 | 0.65 | 3.22 |

# Conclusions & Next Steps

- Majority (>80%) of standardized text entries can be mapped to a standardized concept in the OMOP CDM
- Standardizing how clinical trials data is entered would help support large-scale calculations
- Challenging to evaluate heterogeneity/variance in results

Next Steps:

- Compare with HowOften results
  - Pooled Incidences & Prediction Intervals
  - Variance & Heterogeneity

***Thank you*** *to George Hripcsak, Cindy Chen, Anna Ostropolets, Patrick Ryan, Seung In Seo*