

## 1 Title

2 Constructing a Finer-Grained Representation of Clinical Trial Results from ClinicalTrials.gov

3

## 4 Authors

5 Xuanyu Shi<sup>1,2</sup>, Jian Du<sup>1,2</sup>

6

## 7 Affiliations

8 1. Institute of Medical Technology, Peking University, Beijing, 100191, China

9 2. National Institute of Health Data Science, Peking University, Beijing, 100191, China

10

11 corresponding author: Jian Du ([dujian@bjmu.edu.cn](mailto:dujian@bjmu.edu.cn))

12

## 13 Abstract

14 Randomized controlled trials are essential for evaluating clinical interventions.  
15 ClinicalTrials.gov serves as a primary repository for such data, yet extracting and synthesizing  
16 information from it remains challenging. This study introduces a novel methodology for  
17 constructing a detailed arm-centered representation of clinical trial results, moving beyond  
18 the traditional PICO (Patient, Intervention, Comparison, Outcome) framework. The  
19 representation attentively uncovers both efficacy outcomes and adverse drug events in safety  
20 outcomes, promoting a dual-faceted understanding of intervention effects. Through a  
21 structured acquisition, extraction, and initialization process, we present a knowledge graph  
22 incorporating arm-level efficacy with safety results, categorizing outcomes into three distinct  
23 groups: biomarkers, patient-reported outcomes, and clinical endpoints. The goal is to bridge  
24 the gap between the generally described searchable design information and the specifically  
25 detailed reported results. This approach aims to offer a structured dataset towards better  
26 utilization and interpretation of ClinicalTrials.gov data, facilitating a more feasible and  
27 complete evidence synthesis practice to include both positive and negative results hidden in  
28 clinical trials registries.

29

30

## 31 Background & Summary

32 Clinical trial, specifically randomized controlled trial (RCT), is a method of perspective, random  
33 allocated double-blinded clinical study. The goal of carrying out a clinical trial is to evaluate  
34 the efficacy and safety or a single or multiple clinical interventions. It is also a requisite step  
35 before a drug is approved for market release and sale. In the new evidence-based medicine  
36 pyramid, randomized controlled trials are rated as the highest position, representing the most  
37 reliable medical evidence<sup>1</sup>. In today's medical research landscape, the accuracy and  
38 accessibility of clinical trial data are paramount to ensuring research quality and advancing  
39 scientific progress. ClinicalTrials.gov (CT.gov) serves as a pivotal repository for such data,  
40 hosting valuable information concerning numerous clinical trials. However, extracting and  
41 synthesizing this information for further analysis remains a challenging endeavour.

42

43 The unique value inherent in CT.gov data has been demonstrated through various  
44 comparisons and analyses. When compared with results from PubMed, it has been noted that  
45 CT.gov often contains a more comprehensive report of adverse events<sup>2</sup>. In CT.gov, safety  
46 results were reported at a similar rate as in peer-reviewed literatures, with more thorough  
47 reports of certain safety events<sup>3</sup>.

48

49 Discrepancies between different reports further underscore the necessity for a more  
50 systematic method of information retrieval from CT.gov. Selective reporting, as discussed in  
51 IAMA<sup>4</sup>, illustrates the pressing need for utilizing CT.gov as a tool for evidence synthesis.  
52 Although calls for leveraging CT.gov have been made, including in papers published by BMJ<sup>5</sup>,

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

53 the absence of automated tools for such a task is conspicuous. Despite its seemingly structured  
54 facade, CT.gov presents substantial usability challenges which hinder the extraction of  
55 published results.

56

57 Treating CT.gov research findings as a dataset embodies the principle of "process once, use  
58 many times," transforming these findings into a computable body of knowledge. This  
59 approach stands to significantly aid evidence synthesis. Despite a study from Johns Hopkins  
60 University suggesting that CT.gov has not altered the conclusions of systematic reviews<sup>6</sup>, we  
61 argue that the limited sample size in their analysis may have influenced these findings. Results  
62 of randomized controlled trials (clinical trials, or RCT), are either published in scientific articles,  
63 or reported on registry platforms, such as ClinicalTrials.gov (CT.gov), International Clinical  
64 Trials Registry Platform (ICTRP), and Chinese Clinical Trial Registry, ChiCTR). Researchers found  
65 that in the 91 trials with reported results on ClinicalTrials.gov and published in high-impact  
66 journals, only 52% primary efficacy end points were described in both sources and reported  
67 concordant results<sup>7</sup>. One of the possible and biggest explanation for this phenomenon is  
68 selective publication or publication bias of clinical trials, which means there exists a higher  
69 reporting rate of positive results in published literature compared to that in registries<sup>8</sup>.

70

71 This study intends to construct a finer-grained representation for arm-centered clinical trial  
72 results by integrating efficacy and safety information. Our methodology enables: 1) a detailed  
73 representation of "intervention" entities, transcending the traditional PICO (Patient,  
74 Intervention, Comparison, Outcome) concept representation; 2) a systematic unveiling of  
75 positive and negative results in efficacy outcomes; 3) a systematic disclosure of adverse drug  
76 events (ADE) in safety outcomes; 4) a dual-faceted understanding of intervention effects from  
77 both efficacy and safety perspectives. 5) By providing a structured, ready-to-use dataset, we  
78 aspire to offer a new data source for meta-analysis, thereby facilitating a more discoverable  
79 dataset to enhance evidence-based health decision-making.

80

81 The discrepancy between PICO information in the design phase and reported results  
82 underscores the necessity for real meta-analysis which mandates results-oriented  
83 information. Our study seeks to bridge this gap, fostering a more thorough and reliable  
84 synthesis of evidence that can potentially elevate the standards of medical research and  
85 practice.

86

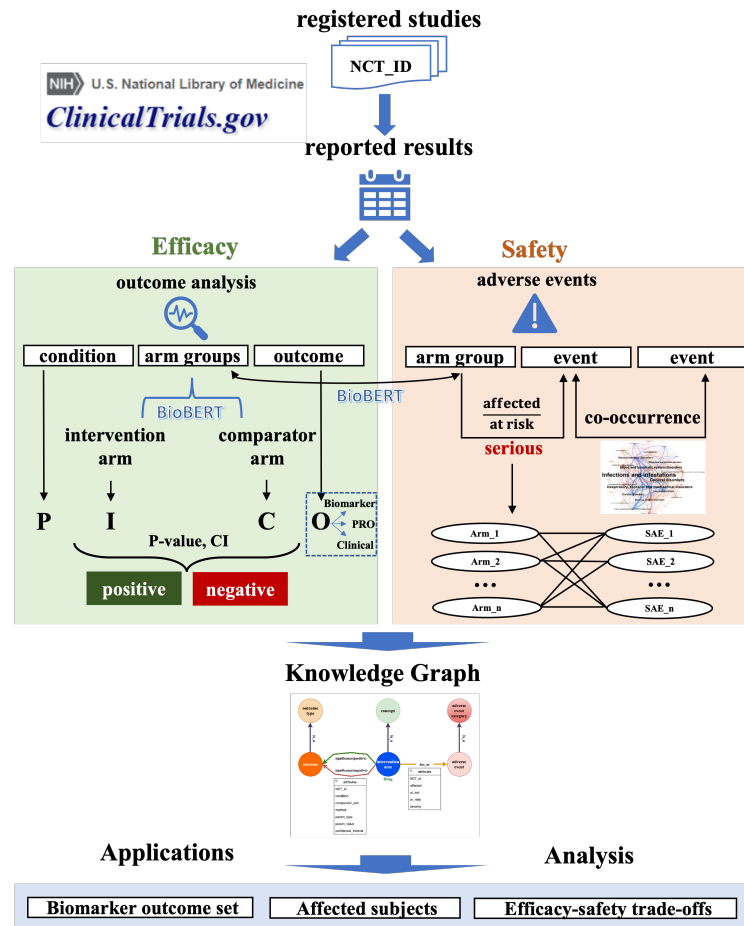
87 **Existing studies.** There have been studies focusing on the structuring trial results data.  
88 MedicineMaps<sup>9</sup> introduced a schema that represents clinical trial results from literature. The  
89 schema lost information about the comparators, and was annotated manually. EvidenceMap<sup>10</sup>  
90 extracted PICO with observation data and efficacy relationships from 80 COVID-19 clinical trial  
91 abstracts. Similarly, entities such as 'two groups' were also labeled as interventional, losing  
92 the ability to build comparable relationships. TrialStreamer<sup>11</sup> and TrialSummarizer<sup>12</sup> are tools  
93 that extracts results from clinical trials in a large scale. For data from registry platform, CTKG<sup>13</sup>  
94 is a large knowledge graph version with embedding analysis of studies in the ClinicalTrials.gov  
95 database. The knowledge graph follows the basic structure of ClinicalTrials.gov, with nodes  
96 being sections and edges being 'has'. In 'outcome\_analysis', same as the representation in  
97 ClinicalTrials.gov, the compared arm groups were not separated as intervention versus  
98 comparator as well.

99

100 **In this study,** we introduce an arm-level representation for efficacy and safety results in  
101 studies registered on ClinicalTrials.gov. We optimize the data structure of clinical trials and  
102 present study results with one-to-one comparable relationships. Each efficacy result is  
103 represented by a relationship from an intervention arm group to the outcome, with the  
104 efficacy as the relationship value, the comparator arm as an attribute. Each safety result is

105 represented by a relationship from an arm group to an adverse event, with the number of  
 106 affected subjects as an attribute.

107  
 108 To represent arm-level efficacy and safety results in a knowledge graph, we started with  
 109 acquiring registered clinical trials from register platform clinicaltrials.gov, followed by  
 110 extraction of statistical significance and adverse events from reported results. After  
 111 standardizing arm groups and outcomes, we eventually constructed a knowledge graph  
 112 including results nodes, relationships and related attributes. Fig. 1 shows the overall workflow  
 113 of this study.



114  
 115 *Figure 1. The overall workflow from data acquisition, results processing, to the final knowledge graph construction*

## 116 **Methods**

117 **Data source.** On December 25<sup>th</sup> in 2022, We downloaded full registered clinical trials dataset  
 118 directly from clinicaltrials.gov (<https://classic.clinicaltrials.gov/AllPublicXML.zip>). The dataset  
 119 is a compressed zip file containing all the individual XML file of each study named by the NCT  
 120 id. The total number of XML files of clinical trials downloaded is 437,173. We parsed and  
 121 transformed each file into a Pandas DataFrame with targeted information. From the reported  
 122 results section, we focus on efficacy and safety related to the intervention arm group, which  
 123 are stored under the statistical analysis and adverse events section.

124  
 125 **Arm-level result.**

126 In the context of clinical trials, an "arm" refers to a group of subjects receiving a particular  
 127 intervention, treatment, or control. The term is often used in randomized controlled trials  
 128 (RCTs), which are considered the gold standard for evaluating the efficacy of different  
 129 treatments. In a typical RCT, participants are randomly assigned to different arms of the study

130 to minimize the influence of confounding variables and to allow for a fair comparison of the  
131 interventions being tested.

132  
133 **Efficacy.** In evidence based medicine, a piece of clinical trial results has to be described as  
134 PICO(Population, Intervention, Comparator, Outcome) framework, along with the efficacy(E).  
135 In clinical trials, population describes the characteristics of the selection of study subjects, such  
136 as age, gender, condition; Intervention represent the treatment or strategy being studied,  
137 which can be a drug, a lifestyle or a dietary plan. Comparator represents the strategy for  
138 comparison with the intervention, which can a placebo method or a standard care. Outcome  
139 describes what is being measure in the study to assess the efficacy and safety of the  
140 intervention, which can be symptom relief, biomedical markers, or mortality rate.

141  
142 The PICO framework is essential for formulating research questions and reviewing literatures  
143 for relevant clinical studies. The search for the efficacy of an intervention, it is essential to  
144 determine the anticipated corresponding population, comparator, and outcome. In natural  
145 language, a clinical trial result can be expressed as: 'There exists a significant difference  
146 between the intervention (I) and the comparator (C) on the outcome (O) in the population (P).  
147

148 In this study, to evaluate the efficacy of an intervention on an outcome, we searched for the  
149 statistical analysis section under each outcome in the study results page. For example, in the  
150 study *NCT01050998*, under the outcome measure of the primary outcome '*Percentage of  
151 Participants Who Achieved Disease Activity Score of 28 Joints Using C-Reactive Protein (DAS28  
152 [CRP]) Response at Day 85 by Region*' exists a list of statistical analysis results. As we can see  
153 from an example of results (Fig. 2), the analysis section provides information such as groups,  
154 p\_value, and confidence interval. With these information we are able to create a relationship  
155 between an intervention arm group and the primary outcome.

156  
157 To parse the XML files for efficacy results, we extracted results in the form of relationships  
158 between 'groups' under each outcome measure and corresponding 'outcome': group--  
159 outcome. Each group—outcome pair is saved in a row in a Pandas dataframe, along with other  
160 columns including other important attributes such as NCT\_ID, p-value, statistical method, and  
161 confidence interval. An example is shown in Table 1.

162

```

<analysis_list>
  <analysis>
    <group_id_list>
      <group_id>01</group_id>
      <group_id>02</group_id>
    </group_id_list>
    <groups_desc>Treatment comparison was made using the multiple imputation
(MI) strategy by combining the results of analysis of covariance (ANCOVA) model with
baseline Hb as a covariate, and treatment, region and cardiovascular/cerebrovascular/
thromboembolic medical history (yes vs. no) as factors.</groups_desc>
    <non_inferiority_type>Non-Inferiority</non_inferiority_type>
    <non_inferiority_desc>Non-inferiority was established if the 2-sided 95%
confidence interval (CI) for the treatment difference in least square (LS) means from
MI ANCOVA model between the 2 treatment groups lay entirely above -0.75 g/dL.</
non_inferiority_desc>
    <p_value>0.0005</p_value>
    <method>ANCOVA</method>
    <method_desc>ANCOVA with multiple imputation</method_desc>
    <param_type>LS Mean Difference</param_type>
    <param_value>0.18</param_value>
    <dispersion_type>Standard Error of the Mean</dispersion_type>
    <dispersion_value>0.053</dispersion_value>
    <ci_percent>95</ci_percent>
    <ci_n_sides>2-Sided</ci_n_sides>
    <ci_lower_limit>0.079</ci_lower_limit>
    <ci_upper_limit>0.287</ci_upper_limit>
  </analysis>
</analysis_list>

```

163  
164

Figure 2. An example of statistical analysis in XML file

165 Table 1. An example of transformed dataframe of efficacy results

NCT_ID	groups	outcome_title	p_value
NCT01049373	['Lymphdiaral Basistropfen (HDC)', 'Placebo Solution']	Change in FFbH-R Between Screening and 2 Weeks	0.0757
NCT03400800	['Inclisiran', 'Placebo']	Percentage Change in LDL-C From Baseline to Day 510	<0.0001
NCT02912650	['Placebo', 'Ibuprofen 250 mg']	Pain Intensity Difference on 11-Point Numerical Scale (PID11)	<0.001

NCT02954354	['Baloxavir', 'Oseltamivir']	Percentage of Participants Reporting Normal Temperature at Each Time Point in Adults Randomized to Baloxavir or Oseltamivir	0.5908
NCT01795547	['Aripiprazole', 'Paliperidone']	Change From Baseline to Week 28 in the 'Interpersonal Relations' QLS Domain Score	0.07

166

167 **Statistical significance.** To better and directly represent the efficacy of an intervention group  
168 on an outcome, we created an automatic rule-based pipeline to summarize analysis attributes  
169 for statistical significance and eventually presented as positive vs negative.

170 (1) check if the result has valid p-value

171 (2) if no p-value found, check if the result has valid confidence interval

172 (3) for results with p-value, we label the results as positive if it has a p-value smaller than  
173 or equal to 0.05, and negative otherwise

174 (4) for results without p-value but with confidence interval, we check whether the  
175 statistical parameter is a ratio-type parameter (Odds Ratio, Hazard Ratio, etc.) or a  
176 difference-type parameter (Mean Difference, Risk Difference ,etc.)

177 (5) for results with a ratio-type statistical parameter, check if the number one (1) is  
178 contained between the lower and upper confidence interval limits. If one is not  
179 contained, label the result as positive, and negative otherwise

180 (6) for results with a difference-type statistical parameter, check if the number zero (0) is  
181 contained between the lower and upper confidence interval limits. If zero is not  
182 contained, label the result as positive, and negative otherwise

183 The detailed visualized pipeline and corresponding number of results in each step are shown  
184 in the result section (Fig. 4).

185

186 In the XML clinical trial files downloaded from the website, arm groups are labeled with  
187 'Experimental' and 'Comparator' in the study design section. But in the statistical analysis  
188 section, the intervention arm group is not distinguished from the comparator arm group by  
189 using a label. The titles of the groups are also not always consistent with the titles from the  
190 study design section, meaning it is not feasible to simply match the labels using the exact  
191 strings. Thus, in this part, we utilized a transformer model BioBERT<sup>14</sup> ('dmis-lab/biobert-v1.1')  
192 to automatically separate intervention arm group from comparator arm group. BERT,  
193 Bidirectional Encoder Representations from Transformers, is a transformers-based deep  
194 learning model that was introduced by Google in 2018<sup>15</sup>. It is a language model that is  
195 pretrained on a large corpus of texts including Wikipedia. The model allows people to invoke  
196 weights (embeddings) of texts from the pretrained model without training on their own.  
197 BioBERT is a domain-specific adaptation of the original BERT and pretrained on a large corpus  
198 on biomedical texts. With BioBERT, we are able to perform various tasks such as named entity  
199 recognition and relation extraction on biomedical problems. In our specific task, we used the  
200 version BioBERT-Base v1.1 that is trained on around one million biomedical publications on  
201 PubMed.

202

203 In this experiment, to prepare the data in the form of the knowledge graph schema, we only  
204 kept results with exact two arm groups for comparison between intervention and comparator.  
205 Studies without valid labels of 'experimental' and 'comparator' arm groups from the study  
206 design section were excluded.

207

208 We first produced semantic embeddings of the arm group titles from both the study design  
209 and statistical analysis sections using BioBERT. Then we calculated the similarities of semantic  
210 embeddings between the arm groups from different sections. Eventually we labeled the arm  
211 group from the statistical analysis section based on the label of the arm group that has the  
212 highest similarity score from the study design section. For examples shown in Table 2, the  
213 'Groups' column are the original stored groups information from ClinicalTrials.gov. After

214 comparing to columns ‘Intervention’ and ‘Comparator’, new columns ‘Intervention\_group’  
215 and ‘Comparator\_group’ were created to represent the experimental and comparator groups.  
216

217 *Table 2 Example results of separation of arm groups into intervention vs. comparator in efficacy results*

NCT_ID	Groups	Intervention	Comparator	Intervention_group	Comparator_group
NCT03064438	[ACU-D1 Ointment, ACU-D1 Ointment Vehicle]	[ACCU-D1]	[Vehicle]	ACU-D1 Ointment	ACU-D1 Ointment Vehicle
NCT02553317	[Caplacizumab, Placebo]	[Caplacizumab]	[Placebo]	Caplacizumab	Placebo
NCT00650806	[Placebo, Canagliflozin 50 mg]	[Canagliflozin (JNJ-28431754)]	[Placebo]	Canagliflozin 50 mg	Placebo
NCT02152605	[Placebo, UMEC/VI 62.5/25 mcg]	[UMEC/VI]	[Placebo]	UMEC/VI 62.5/25 mcg	Placebo
NCT00402246	[Remote Arm, In-Office Arm]	[Remote Management]	[In-Office Care]	Remote Arm	In-Office Arm

218  
219 **Safety.** We integrated reported adverse events from clinicaltrials.gov to represent potential  
220 safety harms related to the interventions. In the trial XML file, an adverse event is stored under  
221 the <event> section, along with <sub\_title> and <count>, recording information including the  
222 corresponding group, the number of affected subjects, and the number of subjects at risk. The  
223 vast majority of adverse event titles are standardized into the Medical Dictionary for  
224 Regulatory Activities (MedDRA<sup>16</sup>) by ClinicalTrials.gov. The parent elements that incorporate  
225 the specific events cover health categories such as ‘Cardiac disorders’, ‘Ear and labyrinth  
226 disorders’, etc. The events are also classified as serious events vs. other events, giving users  
227 the ability to choose by event severity. In this study, we focus on serious adverse events of  
228 which the number of affected subjects is none-zero.

229  
230 In order to intuitively compare the efficacy and safety, we only kept serious adverse events in  
231 the studies that have valid efficacy analysis results (54.8%, 6,429 out of 11,729). On  
232 clinicaltrials.gov, not all the expressions of arm groups in adverse events sections are 100%  
233 same as the ones in efficacy. In order to integrate efficacy and AE, we first tried using string  
234 matching among the arm groups. For arm groups in AE that were not string-matched to an  
235 arm group in efficacy, we utilized BioBert again to calculate the semantic similarities and  
236 return the closest match. For each adverse event title, we calculated the semantic similarities  
237 between the title and all the arm groups in the efficacy section. The closest matched arm group  
238 was saved as a unique column in the dataframe, and we used the matched arm group for  
239 integrating efficacy and safety. The columns for adverse events we saved in the dataframe  
240 include ‘NCT\_ID’, ‘arm\_title’, ‘event\_title’, ‘category’, ‘serious/other’, ‘affected’, ‘at\_risk’, and  
241 ‘matched\_group’. Each row can be comprehended as a reported adverse event named  
242 ‘event\_title’ related to the arm group named ‘matched\_group’ along with the number of  
243 affected and at-risk participants (Table 3).

244  
245 *Table 3. Example adverse events with matched group in efficacy analysis section*

NCT_ID	arm_title	event_title	category	serious/other	affected	at_risk	matched_group
NCT00977080	Oral Paricalcitol in the Oral Stratum	Abdominal pain	Gastrointestinal disorders	serious	1	72	Oral Paricalcitol in the Oral Stratum
NCT02254278	IMRT 5 Weeks	Syncope	Nervous system disorders	serious	1	147	IMRT 5 Weeks
NCT00509145	Laquinimod 0.6 mg	Uterine prolapse	Reproductive system and breast disorders	serious	1	550	Laquinimod
NCT00810199	Tocilizumab + Placebo	Clostridium difficile colitis	Infections and infestations	serious	1	276	Tocilizumab + Placebo
NCT02853331	Sunitinib	Hepatocellular injury	Hepatobiliary disorders	serious	1	425	Sunitinib

246 • **Events co-occurrence.** In extra, we calculated co-occurrence weights among  
247 reported adverse events to provide the which AE is frequently occurred with another  
248 AE. We determined the co-occurrence of categories of SAEs by the following  
249 considerations: if both two categories have reported 'affected' participants in a same  
250 trial, and multiplying the 'affected/at\_risk ratio' from the categories.  
251  $(\sum_{(a,b) \in S, a \neq b} c(a,b) \times (r(a) \times r(b)))$  (c:co-occurrence, r:ratio). This ratio based co-  
252 occurrence took the affected/at\_risk ratio of SAE into account, representing the  
253 probabilistic extent of events co-occurrence.  
254

## 255 Outcome

256 **MeSH extraction.** In clinical trial, expressions of outcomes vary across different studies. Thus  
257 we processed all the unique outcomes using the MeSH(Medical Subject Headings) extraction  
258 tool: NLM Medical Text Indexer(MTI). MTI allows us to extract MeSH terms from natural  
259 languages, in our case, outcomes. MeSH represent the simplest representation of biomedical  
260 languages for tasks including document retrieval, text classification, etc. For all the MeSH  
261 terms extracted, we also recorded the parent terms along with the codings based on the tree  
262 structure of MeSH. With MeSH terms of outcomes, we are able to not only standardize  
263 outcome titles with different expressions, but also classify outcomes into different clinical  
264 categories.  
265

266 **Outcome category.** In clinical trials, the selection of outcomes has critical impact on the  
267 efficacy of the intervention. A study proposed categories of outcomes as Mortality/survival,  
268 Physiological/clinical, Life impact, Resource use, and adverse events<sup>17</sup>. By assumption, with  
269 the same intervention, choosing biomarkers as the outcome potentially tend to produce  
270 positive results<sup>18</sup>. Also, biomarkers are usually used as surrogate endpoints that are easier to  
271 measure. In contrast, in most cases clinical endpoints are more complicate to measure, and  
272 tend to produce negative results. In this study, we classified the outcome titles into three  
273 categories: Biomarker, Patient-reported outcome (PRO), and clinical endpoint.

274 • **Biomarker.** We took the advantage of MeSH to identify biomarker outcomes. We first  
275 manually went through the MeSH tree to discover possible biomarker-related terms.  
276 We found that the majority of the terms we discovered (such as  
277 'E01.370.376.537.250-Brain Cortical Thickness', 'D12.776-Proteins', and 'D10.251-  
278 Fatty Acids') belong to the root category D(Chemicals and Drugs) and E(Analytical,  
279 Diagnostic and Therapeutic Techniques, and Equipments). Thus, we preliminarily  
280 selected the root categories D and E as the biomarker identifier in outcomes. If any of  
281 the MeSH terms extracted from an outcome title starts with D or E, we labeled the  
282 outcome as biomarker. We excluded 'E05.318.308.980-Surveys and Questionnaires'  
283 as it is a PRO related MeSH term. 40,525 results with unique 14,217 biomarker  
284 outcomes were identified.

285 • **PRO.** Patient-reported outcomes(PROs) are used to report the status of a patient's  
286 health condition from the patient's perspective<sup>19</sup>. Research identified PROs and  
287 inspected the inclusion of PROs in registered clinical trials from 2007 to 2013 and  
288 found 27% of the trials used one or more PRO measures<sup>20</sup>. Recently, the usage of PROs  
289 in novel artificial intelligence(AI) clinical trials was assessed by researchers (152/627  
290 trials)<sup>19</sup>. Both previously-mentioned studies used PRO databases such as PROQOLID  
291 and GEM to match PROs in registered clinical trials. One of the major limits of exact  
292 text-matching for outcomes in registered clinical trials is that the writings or  
293 expressions are natural languages, meaning they can be inconsistent. In this study, we  
294 simplified the searching words and took advantage of the extracted MeSH terms. To  
295 identify PRO outcomes, we applied keyword matching on columns 'outcome\_title' and  
296 'outcome\_mesh'. The keywords are 'Survey', 'Questionnaire', 'Patient Reported',

297 'Patient-Reported', 'Scale', and 'Score'. We labeled the outcome as PRO if any  
298 lowercase of the keywords was matched in the outcome string. We eventually found  
299 33,547 results with 9,820 unique PRO outcomes.

300 • **Clinical endpoint.** In this study, we defined the disease-related outcomes as the  
301 clinical endpoint outcomes. The root category C from MeSH tree represents  
302 'Diseases'. We also added F03 (Mental Disorders) as a target disease-related category.  
303 If any of the MeSH terms extracted from an outcome title starts with 'C' and 'F03', we  
304 labeled the outcome as clinical endpoint. 23,557 results with 7,837 unique clinical  
305 endpoint outcomes were identified.

306  
307 In short summary, 25,758 (61.03%) unique outcomes were successfully identified as  
308 biomarker, PRO, or clinical endpoint, contained in 76,575 (63.83%) efficacy results. We noticed  
309 that there existed differences of the coverage of outcome categories between primary  
310 outcomes and secondary outcomes. Table 4 shows the detailed coverage numbers of  
311 biomarkers, PROs, and clinical endpoint outcomes.

312  
313  
314

*Table 4. Coverage statistics of Identified categories in Primary and Secondary outcomes ('result' means numbers in efficacy results, 'outcome' means numbers of unique outcomes in efficacy results)*

	Primary (result)	Secondary (result)	Primary (outcome)	Secondary (outcome)
<b>Total</b>	23,252	92,966	10,624	30,408
Biomarker	9,321 (40.09%)	29,931 (32.20%)	4,105 (38.64%)	9,702 (31.91%)
PRO	4,378 (18.83%)	27,838 (29.94%)	2,072 (19.50%)	7,501 (24.67%)
Clinical	4,133 (17.77%)	18,700 (20.11%)	2,080 (19.58%)	5,537 (18.21%)
<b>coverage</b>	14,388 (61.88%)	59,738 (64.26%)	6,609 (62.21%)	18,429 (60.61%)

315

### 316 Knowledge graph construction.

317 A knowledge graph (KG) is a specialized graph-based data structure used for representing a  
318 collection of knowledge, entities, and their relationships. In this study, we stored the clinical  
319 trial entities (outcome, arm group, adverse event, etc.) and their relationships in a knowledge  
320 graph for visual representation and complex data retrieval queries.

321

322 We created 3 main node types: `intervention_arm`, `outcome`, and `adverse_event`:

323 • **Intervention\_arm.** The `intervention_arm` node is interventional arm group from the  
324 statistical analysis result section. The relationship representing the efficacy of results  
325 is either 'positive' or 'negative', with `intervention_arm` as starting node and `outcome`  
326 as ending node, built based on the columns 'intervention\_group', 'outcome\_title', and  
327 'significance' from each efficacy result row. 'Positive' relationships are shown in green  
328 and 'negative' relationships are shown in red. Some `intervention_arm` nodes are  
329 connected to a concept node that saves the mapped MeSH term of the  
330 `intervention_arm` title.

331 • **Outcome.** The `outcome` node is clinical outcome used in each study. The `outcome`  
332 node can be connected to an `outcome_type` node to exhibit the types of an outcome,  
333 including biomarker, PRO, and clinical endpoint.

334 • **Adverse\_event.** The `adverse_event` node is reported serious adverse event in each  
335 clinical trial. For adverse events, we built the relationship 'has\_ae' from the column  
336 'intervention\_arm' to 'event\_title' from each adverse event result row. Each  
337 `adverse_event` node is connected to an `adverse_event_category` node, representing  
338 the parent biomedical categories of the adverse event name.

339

340 We loaded the created nodes and relationships along with their attributes into the graph  
341 database Neo4j (database version==5.3.0).

342

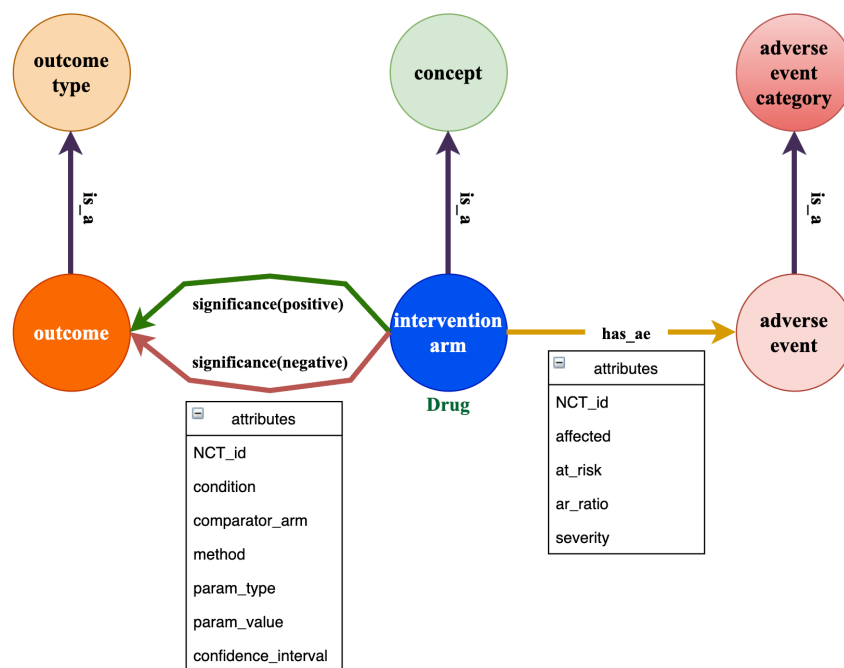


343 To summarize, we presented a pipeline to construct a knowledge graph to represent arm-level  
 344 clinical trial efficacy and safety results. Compared to existing large scale trial databases such  
 345 as AACT and CTKG, we curated a result database containing identified comparable arm groups  
 346 in results section. We also integrated efficacy and safety results by connecting corresponding  
 347 arm groups, providing an evidence dataset for evaluating clinical interventions. We  
 348 constructed a knowledge graph based on the dataset, offering a data infrastructure for further  
 349 analysis and applications. However, current limitations include: (1) In efficacy results, multiple  
 350 testings are not adjusted yet. (2) Multiple arms were not included and classified as  
 351 interventional and comparator. (3) Current dataset only covers clinical trials with both efficacy  
 352 and safety results.

353

## 354 Data Records

355 **Knowledge graph schema.** The knowledge graph schema exhibits the most important nodes,  
 356 relationships, and attributes when evaluating clinical trial results. Compared to the sole ‘has’  
 357 relationship in the CTKG, in order to better represent the RESULTS across different studies, we  
 358 saved all the study-related information of the results in different relationships instead of  
 359 unique nodes, based on the efficacy and adverse event (Fig. 3). For efficacy, the schema shows  
 360 the significance, the comparator arm, the parameter value, NCT\_id, and condition. For safety,  
 361 the schema shows the affected/at\_risk ratio, NCT\_id, and severity. Also, for all the 3 main node  
 362 types, we applied different techniques described in the method section to give each node type  
 363 a standardized or taxonomic categories for data integrity and normalization. Table 5 shows  
 364 the counts of node types and relationships.



365  
 366

Figure 3. An overview of the schema of arm-level efficacy and safety knowledge graph

367 **Table 5. Statistics nodes and efficacy and safety relationships.** The relation count of efficacy is the number of results  
 368 with valid intervention group, comparator group and statistical significance. The relation count of safety is the  
 369 number of serious adverse events with valid arm group and event title.

Node type 0	Count 0	Node type 1	Count 1	Relation name	Relation count
Intervention_arm	10,364	Outcome	41,708	positive/negative	119,968
Intervention_arm	10,364	Adverse_event	40,408	has_ae	803,052
Intervention_arm	10,364	Concept	1439	Is_a	49,764
Outcome	41,708	Outcome type	3	Is_a	76,575
Adverse event	40,408	Adverse event category	28	Is_a	803,052

370



Vaccines	Biological	9,967
Ticagrelor	Drug	9,845
Evolocumab	Biological	9,317
Liraglutide	Drug	9,085

402  
403

Table 7. Numbers of affected subjects by SAE categories

category name	Number of affected subjects
Infections and infestations	294,602
Cardiac disorders	286,035
General disorders	195,018
Gastrointestinal disorders	168,440
Respiratory, thoracic and mediastinal disorders	147,184
Neoplasms benign, malignant and unspecified (incl cysts and polyps)	132,976
Nervous system disorders	129,127
Injury, poisoning and procedural complications	104,784
Vascular disorders	93,766
Blood and lymphatic system disorders	78,164

404

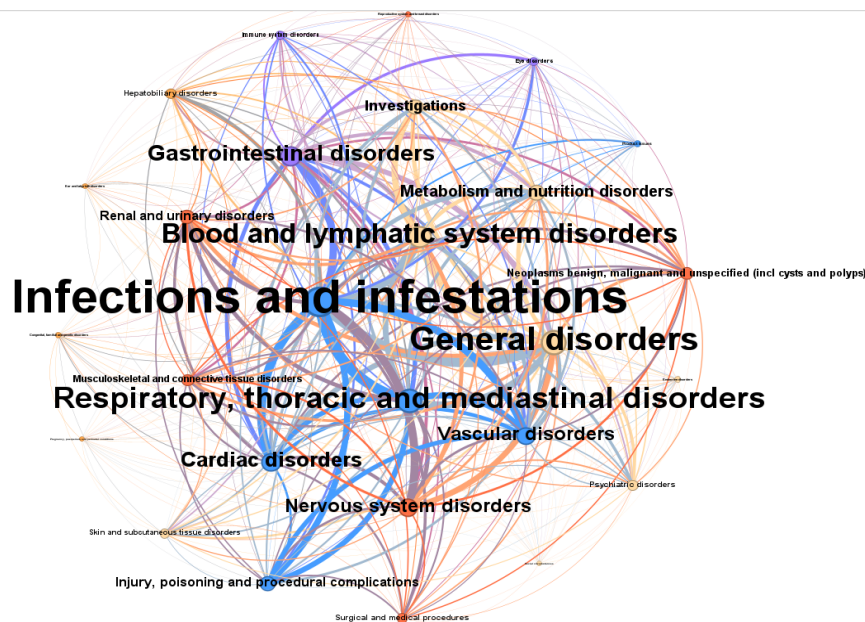


Figure 5 Co-occurrence map of SAE categories

405  
406

407  
408

Table 8. Top co-occurrence edges with ratio weights

SAE category name	SAE category name	Weight
Respiratory, thoracic and mediastinal disorders	Infections and infestations	7.386
General disorders	Infections and infestations	6.818
Cardiac disorders	Infections and infestations	6.003
General disorders	Blood and lymphatic system disorders	5.663
Blood and lymphatic system disorders	Infections and infestations	5.497

409

410

411

412

413

414

415

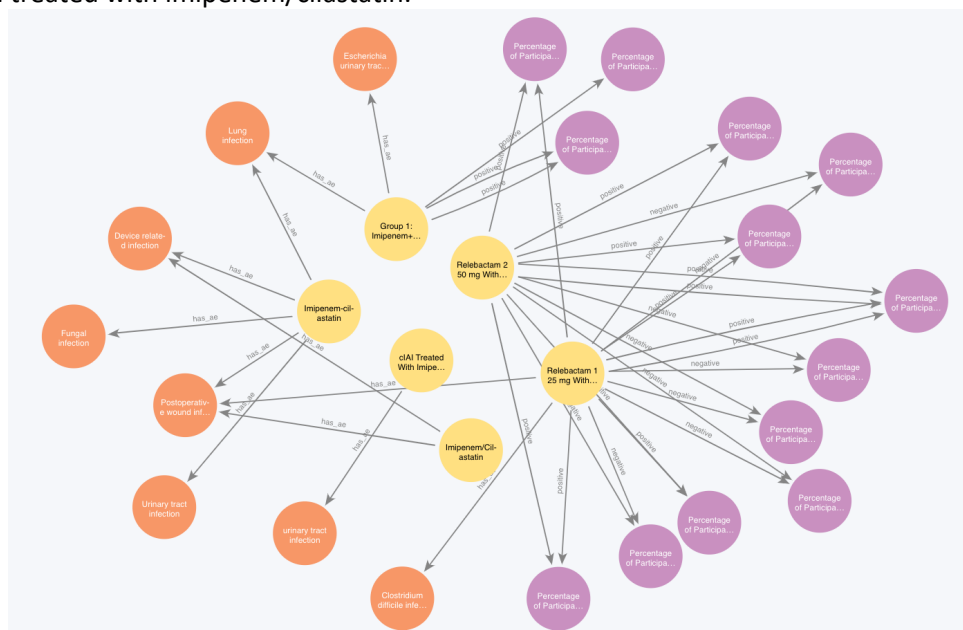
416

417

418

**Knowledge graph: Antibiotic and infection.** In order to present the functionality of evaluating efficacy and safety of an intervention based on the constructed knowledge graph, we were consulted by Na He, a pharmacist from the Department of Pharmacy, Peking University Third Hospital. Na was interested in adverse events related to infections while antibiotic 'imipenem/cilastatin' being used to treat infections. We queried the knowledge graph, retrieved efficacy results by matching the 'intervention\_group' and 'intervention\_MeSH' with 'imipenem' or 'cilastatin', and retrieved adverse event results by matching the group and adverse events title 'event\_title' and parent category 'category' with the keyword 'infection'.

419 The trade-offs between efficacy and safety are crucial in clinical studies, especially the  
 420 population condition and potential adverse events belonging to the same clinical condition.  
 421 The queried relationships are shown in Fig. 6 and detailed SAE results are shown in Table 9.  
 422 The infection-related SAEs give us evidence which types of infections should be paid attention  
 423 when treated with imipenem/cilastatin.



424 *Figure 6. An example KG that shows the case of infection-related adverse events while intervening with the*  
 425 *antibiotic imipenem/cilastatin to treat infections. Orange nodes are SAEs, yellow nodes are interventional arm*  
 426 *groups, purple nodes are outcomes, visualized in Neo4j Bloom*  
 427

428 *Table 9. Infection-related serious adverse events with antibiotic ‘imipenem/cilastatin’*

NCT_ID	arm_title	sub_title	category	affected	at_risk
NCT00515034	cIAI Treated With Imipenem/Cilastatin	urinary tract infection	Infections and infestations	1	19
NCT02321800	Imipenem/Cilastatin	Device related infection	Infections and infestations	1	148
NCT00589693	Imipenem-cilastatin	Urinary tract infection	Infections and infestations	1	112
NCT00589693	Imipenem-cilastatin	Device related infection	Infections and infestations	1	112
NCT00589693	Imipenem-cilastatin	Fungal infection	Infections and infestations	1	112
NCT00589693	Imipenem-cilastatin	Lung infection	Infections and infestations	1	112
NCT00589693	Imipenem-cilastatin	Postoperative wound infection	Infections and infestations	1	112
NCT01506271	Relebactam 125 mg With Imipenem/Cilastatin	Clostridium difficile infection	Infections and infestations	1	116
NCT01506271	Relebactam 125 mg With Imipenem/Cilastatin	Postoperative wound infection	Infections and infestations	1	116
NCT01721408	Imipenem/Cilastatin	Postoperative wound infection	Infections and infestations	2	232
NCT02452047	Group 1: Imipenem+Cilastatin/Relebactam	Lung infection	Infections and infestations	1	31
NCT02452047	Group 3: Open-Label Imipenem+Cilastatin/Relebactam	Escherichia urinary tract infection	Infections and infestations	1	3

429

### 430 Technical Validation

431 We introduced the full pipeline from data collection, data processing to the final knowledge  
 432 graph in details. The knowledge graph and original dataset is transparent in the following

433 aspects: 1) The origin of data comes from ClinicalTrials.gov, which is an open-access platform  
 434 that everyone is able to download the full structured dataset. 2) The process of formation of  
 435 the knowledge graph is reproducible. The codes will be available to access publicly. 3) Users  
 436 are able to adjust the knowledge graph based on provided codes and sub-datasets, providing  
 437 the flexibility to alter the data based on their own research interests. Next, we present three  
 438 case studies to validate the **usability** for exploratory analysis, **consistency** with external  
 439 resources, and **feasibility** for real scientific problems.

440  
 441 **Case study 1: Positive vs. negative efficacy results across types of conditions, interventions,**  
 442 **and outcomes.**

443 We investigated the ratio of positive and negative efficacy results across different dimensions  
 444 (Table 10). For technical validation, we also provide a biomarker outcome list for each  
 445 condition. Each list contains a list of biomarker-type outcomes with corresponding number of  
 446 positive/negative efficacy results. This dataset provides clinical researchers and practitioners  
 447 surrogate endpoints for measurement when it is infeasible, unethical, or ineffective to directly  
 448 evaluate clinical endpoints. This case confirmed the usability of the dataset and users are  
 449 capable of conduct exploratory data analysis from every dimension.

450  
 451

*Table 10. Distribution of positive and negative efficacy results across different dimensions*

	positive	negative	positive_rate	negative_rate
<b><u>Outcome type</u></b>				
Biomarker	14166	15148	48.33%	51.67%
Patient-reported outcome	12647	20900	37.70%	62.30%
Clinical endpoint	9820	13757	41.65%	58.35%
<b><u>Intervention type</u></b>				
Drug	43521	54294	44.49%	55.51%
Biological	5507	7168	43.45%	56.55%
Radiation	57	143	28.50%	71.50%
Behavioral	1019	2375	30.02%	69.98%
Genetic	18	53	25.35%	74.65%
Other	511	966	34.60%	65.40%
Device	1223	1919	38.92%	61.08%
Dietary Supplement	155	389	28.49%	71.51%
Procedure	147	380	27.89%	72.11%
Combination Product	71	48	59.66%	40.34%
Diagnostic Test	1	3	25.00%	75.00%
<b><u>Health Problems</u></b>				
Infections	2640	3660	41.90%	58.10%
Neoplasms	1152	1929	37.39%	62.61%
Musculoskeletal Diseases	4338	5108	45.92%	54.08%
Digestive System Diseases	1007	1352	42.69%	57.31%
Stomatognathic Diseases	544	660	45.18%	54.82%
Respiratory Tract Diseases	4520	6337	41.63%	58.37%
Otorhinolaryngologic Diseases	520	644	44.67%	55.33%
Nervous System Diseases	2090	5086	29.12%	70.88%
Eye Diseases	478	940	33.71%	66.29%
Urogenital Diseases	1438	1461	49.60%	50.40%
Female Urogenital Diseases and Pregnancy Complications	1614	1621	49.89%	50.11%
Cardiovascular Diseases	2071	2647	43.90%	56.10%
Hemic and Lymphatic Diseases	434	639	40.45%	59.55%
Congenital, Hereditary, and Neonatal Diseases and Abnormalities	1398	1325	51.34%	48.66%
Skin and Connective Tissue Diseases	3539	3870	47.77%	52.23%
Nutritional and Metabolic Diseases	6091	5331	53.33%	46.67%
Endocrine System Diseases	3546	3190	52.64%	47.36%
Immune System Diseases	4425	6122	41.96%	58.04%

Pathological Conditions, Signs and Symptoms	5483	7817	41.23%	58.77%
Occupational Diseases	8	174	4.40%	95.60%
Chemically-Induced Disorders	107	150	41.63%	58.37%
Wounds and Injuries	131	331	28.35%	71.65%
Mental Disorders	2408	4575	34.48%	65.52%

452

453 **Biomarker outcome set.** We provide a dataset of biomarker outcomes with the most positive  
454 and negative efficacy results under each health problem. The dataset is provided in  
455 Supplementary Table 1. Note that the outcomes are not standardized by concept, meaning  
456 there might be variations of same outcomes that are not summed together. Besides, this  
457 example shows the top biomarker outcomes associated with health problems. In the full  
458 dataset, outcomes should be associated with the original condition name.

459

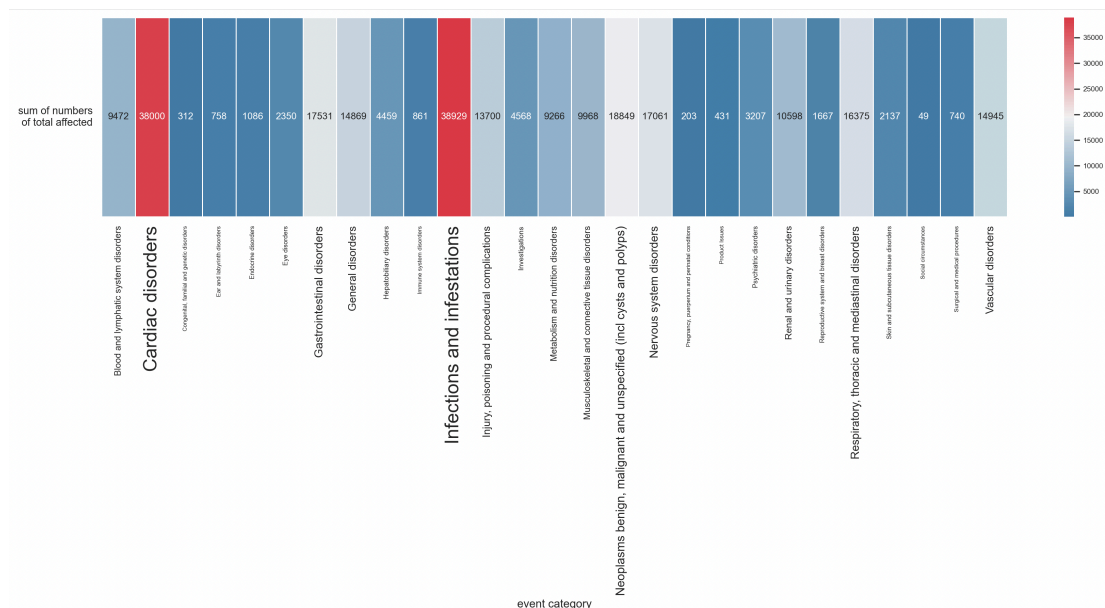
460 **Case study 2: Top SAE categories comparison with FDA data**

461 US Food and Drug Administration (FDA) manages an adverse event reporting database (FDA's  
462 Adverse Event Reporting System, FAERS) intending to support post-marketing safety  
463 surveillance for drug and biomedical products. Compared to RCT results on CT.gov, FAERS  
464 receive reports directly from healthcare professionals and consumers.

465

466 Simon Lafosse visualized the distribution of adverse events by body systems from FAERS<sup>22</sup>. The  
467 visualization shows that systems such as 'General Disorders', 'Injury, Poisoning', and 'Skin'  
468 have most of the reported AE cases from April 2018 to March 2019. Inspired of this work, with  
469 results data of ClinicalTrials.gov in hand, we inspected the sum of numbers of affected subjects  
470 under each SAE category between 2018 and 2019 (Fig. 7). In this case, we included all the  
471 possible interventions (top 20 in Simon Lafosse's), and also visualized the distribution of SAE  
472 by categories. 'Infections and infestations' and 'Cardiac disorders' are related to 38,929 and  
473 38,000 reported affect subjects in this dataset. This case validates that our dataset can be used  
474 to analyze SAE in clinical trials on a large scale. Also compared to the analysis of FDA data, we  
475 found some common top condition categories such as 'Nervous system disorders', and  
476 'Infections and infestations'. There is great potential of deeper analysis of SAE based on this  
477 dataset since it has multiple levels of meta-data including co-occurrence weights, mapped  
478 categories, corresponding efficacy results, etc. This case shows the consistency of ability to  
479 analyze adverse events by condition category from ClinicalTrials.gov data, similar to other  
480 analysis based on FDA data.

481

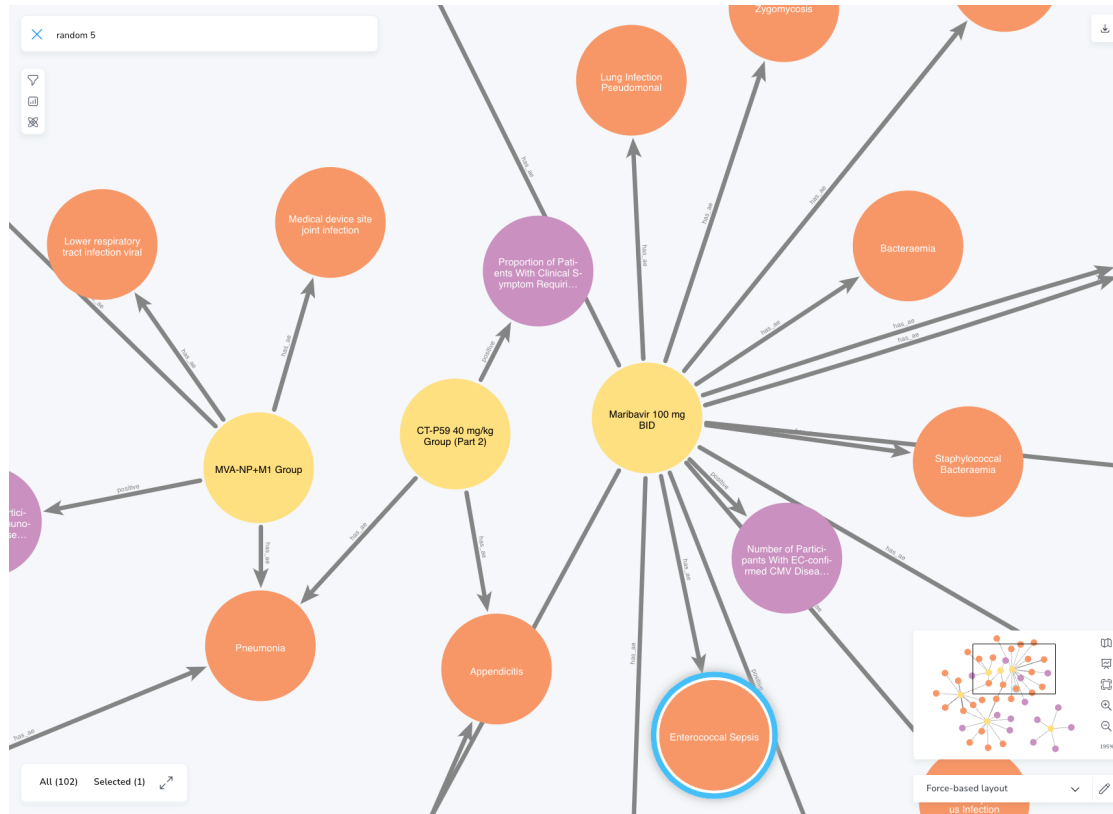


482  
483

Figure 7. The distribution of health problems of serious adverse events by the number of affected subjects

484 **Case study 3: Interventions with same condition and SAE category.**

485 Multimorbidity and comorbidity are vital challenges to medical practice<sup>23,24</sup>. We attempt to  
 486 find interventional arm groups found having associated SAE categories while used to treat the  
 487 same category. In this case, we selected ‘Infections and Infestations’ as the category since it is  
 488 the category that is associated most trial subject. Fig. 8 illustrates the KG retrieval of above  
 489 circumstance, with yellow nodes are intervention arm groups, orange nodes are SAEs, and  
 490 purple nodes are outcomes. This case validates that with real medical practice problem, this  
 491 knowledge graph has much potential to provide reliable, large-scale, and human-readable  
 492 evidence.



493  
 494 *Figure 8. KG example in which both efficacy and safety outcomes are infections, visualized in Neo4j Bloom*

495 We retrieved the arm groups and corresponding mesh terms and saved them to provide an  
 496 intervention set for researchers who want to find interventions that are tested positive for  
 497 efficacy and are observed with SAEs from the same condition category at the same time. This  
 498 knowledge graph can also be used to retrieve for other conditions. The full arm groups set and  
 499 corresponding mesh terms are in the following:

500  
 501 **Arm groups set:** ['Vaniprevir 24 Week Arm', 'Rabipur Group', '7vPnC After Infant Series Dose  
 502 3', 'MK-3415 + SOC', 'Nevirapine QD+BiD', 'Vaccine', 'V710 60 µg', 'CD24Fc', 'Enhanced  
 503 Treatment', 'Priorix-Tetra Group', 'CCI Vaccine', 'Group D: TDV Lyophilized', 'Control',  
 504 'HIV+/Cervarix Group', 'rhNGF 20 µg/ml', 'Telaprevir 24 Week+Peg-IFN-alfa-2a,RBV 48 Week',  
 505 'V419', 'Placebo + INFANRIX Hexa', 'QIVc (≥4 to <18 Years)', 'Baricitinib + SOC', 'ADV-TDF',  
 506 'Ziprasidone', 'Reltecimod (AB103) 0.5 mg/kg', 'Cohort 1-Engerix-B/Prevnar-Hiberix Group',  
 507 'Arm 1: Metronidazole Plus Miconazole', 'Neutrolin Arm', 'SB213503 Lot 1 + M-M-R Group',  
 508 'Azithromycin', 'DTG + RPV', 'Twinrix Group', 'IVV Vaccine', 'QIVc (≥18 Years)', '7vPnC After the  
 509 Infant Series', 'ABC/ DTG/ 3TC', 'Immediate Introduction of Everolimus', 'Baloxavir Marboxil',  
 510 'PTX+ART', 'RotaTeq™', '4200mg Bamlanivimab', 'MMR Group', 'DTaP (Catch-up 7vPnC)', 'CAB  
 511 LA+RPV LA (Q4W)', 'Placebo - Asia', 'Immediate ART', 'Telaprevir 12 Week+Peg-IFN-alfa-  
 512 2a,RBV 48 Week', 'Letermovir', 'Cyclosporine', 'Bimekizumab (PPS)', 'Maribavir 400 mg', 'V501',  
 513 'Investigational Device', 'Oral Care With Influenza and Pneumococcal Vaccines', 'MK-6072 +

514 SOC', 'Placebo for 24 Weeks', 'TMC435 150mg 12Wks PR24/48', 'Cohort 2', 'Td Adsorbed  
515 Vaccine', 'Baloxavir', 'Pneumosil', 'CT-P59 40 mg/kg Group (Part 2)', 'Test', 'Delafloxacin',  
516 'Zoster-022 GSK1437173A 70-79YOA Group', 'Group 1: QIV-HD', 'Phase 3: 700 mg  
517 Bamlanivimab + 1400 mg Etesevimab', 'Cohort 3', 'Azithromycin + Chloroquine', 'Maribavir 100  
518 mg BID', 'Camostat Mesilate', '7vPnC After 2-Dose Infant Series', 'V260 With Staggered or  
519 Concomitant EPI', 'Anaferon for Children', 'M72AS01 Group', 'Double-Blind Period - Vicriviroc  
520 20 mg Plus an ART Regimen', 'MEDI8897 50 mg', 'MK-8237', 'TIV (Elderly)', 'Peginterferon Alfa-  
521 2a + Ribavirin', 'Clinically Driven Monitoring (CDM)', 'Telaprevir 12 Week +Peg-IFN-alfa-2a, RBV  
522 24 Week', 'Raltegravir 800 mg q.d.', 'V419 Lots A, B, and C Combined', 'MK-3415A + SOC',  
523 'Combined Patients Receiving 74 Gy RT', 'SYN-004', 'Nevirapine', 'Cohort 3 (Placebo, 2-5 Yrs)',  
524 'Camostat Mesylate', 'Telaprevir 8 Week, PBO 4 Week+Peg-IFN-alfa-2a, RBV 24/48 Week',  
525 'DRV/COBI+ FTC/TDF (Control) (Baseline to Switch)', 'TDV Lot 3', 'qHPV Vaccine in Base Study',  
526 'Telaprevir 12 Week+Peg-IFN-alfa-2a, RBV 24 Week', 'Double-Blind Period - Vicriviroc 30 mg  
527 Plus an ART Regimen', 'MVA-NP+M1 Group', 'Maternal Chloroquine Prophylaxis', 'FluLaval™  
528 Quadrivalent Group', '4\_rhNGF10\_Phase 2\_treatment', 'V114', 'High-Dose Quadrivalent  
529 Influenza Vaccine (QIV-HD)', 'GSK1437173A Group', 'PR5I', 'Stopped Cotrimoxazole  
530 Prophylaxis', 'GSK1437173A Overall Ages Group', 'AZD7442', 'Telaprevir 24 Week+Peg-IFN-  
531 alfa-2a 24 Week', 'Test Product', 'TOTAL 2-sites, TRC', 'Cohort 1', 'Arm A: hIVIG', 'Posaconazole  
532 100 mg QD for 24 Weeks', 'Remdesivir', 'Menjugate Comparator [6-12W] Group', 'HIV-  
533 /Cervarix Group', 'Vaniprevir 12 Week Arm', 'Participants Who Received DTG/3TC FDC', '4-sites,  
534 1-week WITH HRIG', 'Daclatasvir, 60 mg, 12-Week Cohort', 'Intervention', '13vPnC', 'Arm B:  
535 ZDV+ABC+3TC+NNRTI->ABC+3TC+NNRTI Maintenance', 'VeroRab Comparator [5-17M] Group',  
536 'Group D', 'PEG 1.0 mcg/kg Weekly (QW) \* 24 Weeks', 'Cohort 3 eTIV (3-8 Years)', 'Sotrovimab  
537 (500 mg IV)', 'Filgotinib', 'rMenB All', 'Azelaic Acid Foam, 15% (BAY39-6251)', 'QIV-HD',  
538 'Placebo', 'TIV (3-8 Years)', 'Phase 2: 2800 mg Bamlanivimab + 2800 mg Etesevimab', 'TDV Lot  
539 2', '13vPnC After Toddler Dose', 'AZD1222', 'Telaprevir 12 Week+Peg-IFN-alfa-2a, RBV 24/48  
540 Week', 'Agriflu', 'Period 2: Tacrolimus']

541

542 **Mesh terms set:** ['taurolidine', 'peginterferon alfa-2b', 'metronidazole', 'chloroquine',  
543 'paclitaxel', 'acetaminophen', 'valganciclovir', 'lamivudine', 'efavirenz, emtricitabine, tenofovir  
544 disoproxil fumarate drug combination', 'terbinafine', 'loratadine', 'vincristine', 'adapalene',  
545 'sulfadoxine', 'everolimus', 'antibodies', 'interferon alpha-2', 'ribavirin', 'trimethoprim,  
546 sulfamethoxazole drug combination', 'rho(d) immune globulin', 'immunoglobulins',  
547 'methylprednisolone hemisuccinate', 'thrombin', 'peginterferon alfa-2a', 'darunavir',  
548 'mycophenolic acid', 'alemtuzumab', 'diphenhydramine', 'nevirapine', 'tenofovir', 'cyclosporine',  
549 'infliximab', 'cetuximab', 'cyclosporins', 'sirolimus', 'cobicistat', 'promethazine', 'gamma-  
550 globulins', 'trimethoprim', 'mometasone furoate', 'tofacitinib', 'emtricitabine, tenofovir  
551 disoproxil fumarate drug combination', 'thymoglobulin', 'maribavir', 'carboplatin',  
552 'bamlanivimab', 'basiliximab', 'pharmaceutical solutions', 'gelatin sponge, absorbable',  
553 'heparin', 'atazanavir sulfate', 'azelaic acid', 'fanasil, pyrimethamine drug combination',  
554 'calcium heparin', 'clindamycin', 'tazarotene', 'vaccines', 'posaconazole', 'interferons',  
555 'olopatadine hydrochloride', 'heptavalent pneumococcal conjugate vaccine', 'ceftriaxone',  
556 'ivermectin', 'simeprevir', 'rilpivirine', 'clotrimazole', 'clindamycin phosphate', 'ganciclovir',  
557 'sulfamethoxazole', 'oseltamivir', 'abacavir', 'raltegravir potassium', 'miconazole', 'mitogens',  
558 'adalimumab', 'amoxicillin', 'tacrolimus', 'nicotinic acids', 'antibodies, monoclonal',  
559 'epinephrine', 'interferon-alpha', 'remdesivir', 'pyrimethamine', 'cobicistat mixture with  
560 darunavir', 'camostat', 'ziprasidone', 'efavirenz', 'immunoglobulins, intravenous', 'artemether,  
561 lumefantrine drug combination', 'adefovir dipivoxil', 'gabexate', 'clindamycin palmitate',  
562 'letermovir', 'methylprednisolone', 'ganciclovir triphosphate', 'emtricitabine tenofovir  
563 alafenamide', 'ophthalmic solutions', 'zidovudine', 'lamivudine, zidovudine drug combination',  
564 'prednisone', 'emtricitabine', 'amodiaquine', 'benzoyl peroxide', 'azithromycin', 'etoposide',  
565 'antilymphocyte serum', 'bleomycin', 'dideoxynucleosides']



## 566 **Code Availability**

567 The source codes of data collection, processing and analysis are stored at:  
568 (<https://github.com/xuanyshi/clinical-trials-finer-grained-representation>).

569

## 570 **Acknowledgements**

571 This work is funded by the National Key R&D Program for Young Scientists (2022YFF0712000).

572

573 We thank Na He (pharmacist, Department of Pharmacy, Peking University Third Hospital) for  
574 drug efficacy and adverse events scenarios. We also thank research interns Xiaofan Li (Beijing  
575 Institute of Technology), Zitao Liang (Peking University), and Zhanyuan Jiang (Beijing University  
576 of Posts and Telecommunications) for help with data processing and analysis.

577

## 578 **Author contributions**

579 Xuanyu Shi was responsible for major data collection, pre-processing and analysis, XS was also  
580 responsible for manuscript writing.

581 Jian Du and Xuanyu Shi conceptualized the research. JD was also responsible for background  
582 drafting and manuscript polishing.

583

## 584 **Competing interests**

585 We declare no conflict of interest.

586

## 587 **References**

- 588 1 Murad, M. H., Asi, N., Alsawas, M. & Alahdab, F. New evidence pyramid. *BMJ Evidence-*  
589 *Based Medicine* **21**, 125-127 (2016).
- 590 2 Dunn, A. G. & Bourgeois, F. T. Is it time for computable evidence synthesis? *Journal of*  
591 *the American Medical Informatics Association* **27**, 972-975 (2020).
- 592 3 Chen, K. Y., Borglund, E. M., Postema, E. C., Dunn, A. G. & Bourgeois, F. T. Reporting  
593 of clinical trial safety results in ClinicalTrials.gov for FDA-approved drugs: A cross-  
594 sectional analysis. *Clinical Trials* **19**, 442-451 (2022).
- 595 4 Chan, A.-W., Hróbjartsson, A., Haahr, M. T., Gøtzsche, P. C. & Altman, D. G. Empirical  
596 evidence for selective reporting of outcomes in randomized trials: comparison of  
597 protocols to published articles. *Jama* **291**, 2457-2465 (2004).
- 598 5 Jia, Y. *et al.* Effect of redundant clinical trials from mainland China evaluating statins  
599 in patients with coronary artery disease: cross sectional study. *bmj* **372** (2021).
- 600 6 Wilson, L. M., Sharma, R., Dy, S. M., Waldfoegel, J. M. & Robinson, K. A. Searching  
601 ClinicalTrials.gov did not change the conclusions of a systematic review. *Journal of*  
602 *Clinical Epidemiology* **90**, 127-135 (2017).
- 603 7 Becker, J. E., Krumholz, H. M., Ben-Josef, G. & Ross, J. S. Reporting of results in  
604 ClinicalTrials.gov and high-impact journals. *Jama* **311**, 1063-1065 (2014).
- 605 8 Turner, E. H., Matthews, A. M., Linardatos, E., Tell, R. A. & Rosenthal, R. Selective  
606 publication of antidepressant trials and its influence on apparent efficacy. *New*  
607 *England Journal of Medicine* **358**, 252-260 (2008).
- 608 9 Mainberger, F. *et al.* MedicineMaps: A tool for mapping and linking evidence from  
609 experimental and clinical trial literature. (2013).
- 610 10 Kang, T. *et al.* EvidenceMap: a three-level knowledge representation for medical  
611 evidence computation and comprehension. *Journal of the American Medical*  
612 *Informatics Association* **30**, 1022-1031 (2023).
- 613 11 Marshall, I. J. *et al.* Trialstreamer: A living, automatically updated database of clinical  
614 trial reports. *Journal of the American Medical Informatics Association* **27**, 1903-1912  
615 (2020).

- 616 12 Ramprasad, S., Marshall, I. J., McInerney, D. J. & Wallace, B. C. in *Proceedings of the*  
617 *conference. Association for Computational Linguistics. Meeting.* 236 (NIH Public  
618 Access).
- 619 13 Chen, Z. *et al.* A knowledge graph of clinical trials (CTKG). *Scientific reports* **12**, 4724  
620 (2022).
- 621 14 Lee, J. *et al.* BioBERT: a pre-trained biomedical language representation model for  
622 biomedical text mining. *Bioinformatics* **36**, 1234-1240 (2020).
- 623 15 Devlin, J., Chang, M.-W., Lee, K. & Toutanova, K. Bert: Pre-training of deep  
624 bidirectional transformers for language understanding. *arXiv preprint*  
625 *arXiv:1810.04805* (2018).
- 626 16 Brown, E. G., Wood, L. & Wood, S. The medical dictionary for regulatory activities  
627 (MedDRA). *Drug safety* **20**, 109-117 (1999).
- 628 17 Dodd, S. *et al.* A taxonomy has been developed for outcomes in medical research to  
629 help improve knowledge discovery. *Journal of clinical epidemiology* **96**, 84-92 (2018).
- 630 18 Fleming, T. R. & Powers, J. H. Biomarkers and surrogate endpoints in clinical trials.  
631 *Statistics in medicine* **31**, 2973-2984 (2012).
- 632 19 Pearce, F. J. *et al.* The role of patient-reported outcome measures in trials of artificial  
633 intelligence health technologies: a systematic evaluation of ClinicalTrials. gov records  
634 (1997–2022). *The Lancet Digital Health* **5**, e160-e167 (2023).
- 635 20 Vodicka, E. *et al.* Inclusion of patient-reported outcome measures in registered clinical  
636 trials: evidence from ClinicalTrials. gov (2007–2013). *Contemporary clinical trials* **43**,  
637 1-9 (2015).
- 638 21 Bastian, M., Heymann, S. & Jacomy, M. in *Proceedings of the international AAAI*  
639 *conference on web and social media.* 361-362.
- 640 22 Lafosse, S. *Watch your medicine drugs: the adverse events and medicinal drugs,*  
641 [https://public.tableau.com/app/profile/simon.lafosse/viz/WYMD-](https://public.tableau.com/app/profile/simon.lafosse/viz/WYMD-Top20v0_5/WYMD)  
642 [Top20v0\\_5/WYMD](https://public.tableau.com/app/profile/simon.lafosse/viz/WYMD-Top20v0_5/WYMD)> (2019).
- 643 23 Kuan, V. *et al.* Identifying and visualising multimorbidity and comorbidity patterns in  
644 patients in the English National Health Service: a population-based study. *The Lancet*  
645 *Digital Health* **5**, e16-e27 (2023).
- 646 24 Guthrie, B., Payne, K., Alderson, P., McMurdo, M. E. & Mercer, S. W. Adapting clinical  
647 guidelines to take account of multimorbidity. *Bmj* **345**, e6341 (2012).

648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663

664  
665  
666

Supplementary Table 1. Positive and negative biomarker outcome set of each health problem

health_problem_title	health_problem_title	outcome_title	positive_number	outcome_title	negative_number
C01	Infections	Geometric Mean Titer of Serotype-specific Oposonophagocytic Activity (OPA) Following Vaccination With Separate V114 Lots	45	Percentage of Participant With Normal Alanine Aminotransferase (ALT) Levels	40
C01	Infections	Percentage of Participants Achieving the Serotype-specific Pneumococcal Immunoglobulin G (IgG) Antibody Threshold Value of $\geq 0.35$ $\mu\text{g}/\text{mL}$ for the 13 Common Serotypes in V114 and Pevnar 13™: 1 Month Post Vaccination 3	26	Percentage of Participants Who Were HBeAg Negative	33
C01	Infections	Ratio to Baseline in Renal Biomarkers-Urine and Serum Beta-2 Microglobulin (B2M), Urine Albumin/Creatinine, Urine B2M/Urine Creatinine, Urine Phosphate, Urine Protein/Creatinine, Urine RBP 4 and Urine RBP 4/Urine Creatinine at Weeks 24, 48	22	Change From Baseline in HBsAg Levels at Weeks 4, 8, 16, 28, 40, 52, 64, 76, 88, and 100	31
C01	Infections	Serotype-specific Oposonophagocytic Activity (OPA) Geometric Mean Titers (GMTs) 1 Month After Vaccination for 12 Common Serotypes in 13vPnC/23vPS Group Relative to 23vPS Group and 23vPS/23vPS Group	21	Serotype-specific Geometric Mean Concentration of IgG Antibody	21
C01	Infections	Comparison of Serotype-specific Geometric Mean Concentration of IgG Antibody Response 4 Weeks After a 3-dose Primary Series to 4 Weeks After a Booster Dose	20	Change From Baseline in Hepatitis B Virus Deoxyribonucleic Acid (HBV DNA) Levels at Weeks 4, 8, 16, 28, 40, 52, 64, 76, 88, and 100	19
C04	Neoplasms	Progression-Free Survival (PFS)	40	PFS in Subgroups That Were Defined by Germline PDGFRB Polymorphisms	48
C04	Neoplasms	AUC(0-2h) of [6R]-5,10-methylene-THF	6	OS in Subgroups That Were Defined by Germline PDGFRB Polymorphisms	46
C04	Neoplasms	AUC(0-2h) of [6S]-5-THF	6	PFS in Subgroups That Were Defined by RNA Expression Profile	35
C04	Neoplasms	AUC(0-2h) of [6SR]-5-formyl-THF	6	Overall Survival (OS) in Subgroups That Were Defined by Germline VEGFR2 Polymorphisms	24
C04	Neoplasms	AUC(Last) of [6R]-5,10-methylene-THF	6	PFS in Subgroups That Were Defined by Germline VEGFR2 Polymorphisms	22
C05	Musculoskeletal Diseases	Change From Baseline in Serum Uric Acid Levels at Day 1, Day 3, Day 7, Day 11, Day 14 and Follow-up	44	Change From Baseline in Plasma Levels of Hypoxanthine at Day 1, Day 7, Day 14, and at Follow-up	36
C05	Musculoskeletal Diseases	Change From Baseline in Erythrocyte Sedimentation Rate (ESR) at Weeks 2, 4, 8, 12, 16, 20, 24, 32, 40, 48, and 52	27	Current Pain During the Hospital Stay Assessed by the Pain NRS	26
C05	Musculoskeletal Diseases	Change From Baseline in C-reactive Protein (CRP) at Weeks 2, 4, 8, 12, 16, 20, 24, 32, 40, 48, and 52	24	Change From Baseline in Predictive Biomarkers: Amyloid A, Chemokine (C-C Motif) Ligand 17, Chemokine (C-X-C Motif) Ligand 13, Interleukin 6, Macrophage-Derived Chemokine	26
C05	Musculoskeletal Diseases	Change From Baseline in Plasma Levels of Xanthine at Day 1, Day 7, Day 14, and at Follow-up	22	Change From Baseline in Plasma Levels of Xanthine at Day 1, Day 7, Day 14, and at Follow-up	20
C05	Musculoskeletal Diseases	Percent Change From Baseline in Serum Urate Levels at Week 28.	13	Change From Baseline in C-Reactive Protein (CRP) at Weeks 1, 2, 4, 8, 12 and 16	18
C06	Digestive System Diseases	Change From Baseline in Quantitative Hepatitis B Surface Antigen (Log qHBsAg) Over Time	17	Percentage of Participant With Normal Alanine Aminotransferase (ALT) Levels	40
C06	Digestive System Diseases	Change From Baseline in Hepatitis B Virus Deoxyribonucleic Acid (HBV DNA) Levels at Weeks 4, 8, 16, 28, 40, 52, 64, 76, 88, and 100	14	Percentage of Participants Who Were HBeAg Negative	33
C06	Digestive System Diseases	Absolute Change in Percent Predicted Forced Expiratory Volume in 1 Second (ppFEV1)	10	Change From Baseline in HBsAg Levels at Weeks 4, 8, 16, 28, 40, 52, 64, 76, 88, and 100	31
C06	Digestive System Diseases	Progression Free Survival (PFS)	6	Change From Baseline in Hepatitis B Virus Deoxyribonucleic Acid (HBV DNA) Levels at Weeks 4, 8, 16, 28, 40, 52, 64, 76, 88, and 100	19
C06	Digestive System Diseases	AUC(0-2h) of [6R]-5,10-methylene-THF	6	Progression Free Survival (PFS)	6
C07	Stomatognathic Diseases	AUCgly(0-90) for Test Zinc-IPMP, Test Zinc Non-IPMP, Positive Control, Non-SLS Negative Control and SLS Negative Control	6	Anti-S.Pneumoniae Antibody Concentration in US Sub-cohort of Pooled MMR Groups	12
C07	Stomatognathic Diseases	Enamel Fluoride Uptake	5	AUClive:Dead(0-90) for Test Zinc-IPMP, Test Zinc Non-IPMP, Positive Control, Non-SLS Negative Control and SLS Negative Control	10
C07	Stomatognathic Diseases	AUCregrowth(0-90) for Test Zinc-IPMP, Test Zinc Non-IPMP, Positive Control, Non-SLS Negative Control and SLS Negative Control	4	AUCregrowth(0-90) for Test Zinc-IPMP, Test Zinc Non-IPMP, Positive Control, Non-SLS Negative Control and SLS Negative Control	6
C07	Stomatognathic Diseases	Adjusted Mean Percent Net Erosion Resistance (NER) of Enamel Specimens Exposed to Test Dentifrice + Test MR Relative to: 1) Test Dentifrice+ Sterile Water Rinse 2) Reference Dentifrice+ Sterile Water Rinse 3) Placebo Dentifrice+ Sterile Water Rinse	3	AUCgly(0-90) for Test Zinc-IPMP, Test Zinc Non-IPMP, Positive Control, Non-SLS Negative Control and SLS Negative Control	3
C07	Stomatognathic Diseases	Adjusted Mean Percentage Surface Microhardness (SMH) Recovery of Enamel Specimens Exposed to Test Dentifrice + Test MR Relative to: 1)Test Dentifrice+Sterile Water Rinse 2)Reference Dentifrice+Sterile Water Rinse 3)Placebo Dentifrice+ Sterile Water Rinse	3	Number of Participants With Unplanned Breaks in Cisplatin Chemotherapy Treatment	2
C08	Respiratory Tract Diseases	Geometric Mean Titer of Serotype-specific Oposonophagocytic Activity (OPA) Following Vaccination With Separate V114 Lots	45	Change From Baseline in Specific Imaging Airway Volume (siVaw), Measured at FRC and TLC Scan Conditions, Presented in Longitudinal and Scan Trimmed Scan Types, Measured in 5 Lobes and 5 Regions at Screening, Day 12 and Day 28	100
C08	Respiratory Tract Diseases	Change From Baseline in Fraction of Exhaled Nitric Oxide (FeNO) Over Time Following the Cessation of Repeat Dose Treatment With FF/VI	34	Change From Baseline in Imaging Airways Volume: iVaw, Measured at FRC and TLC Scan Conditions, Presented in Longitudinal and Scan Trimmed Scan Types, Measured in 5 Lobes and 5 Regions at Screening, Day 12 and Day 28	100
C08	Respiratory Tract Diseases	Change From Baseline in Morning and Evening Peak Expiratory Flow Rate (PEF) Over 26 and 52 Weeks of Treatment	28	Change From Baseline in Imaging Airways Resistance (iRaw) Measured at FRC and TLC Scan Conditions, Presented in Scan Trimmed Scan Types, Measured in 5 Lobes and 5 Regions at Screening, Day 12 and Day 28	60
C08	Respiratory Tract Diseases	Comparison of Serotype-specific Geometric Mean Concentration of IgG Antibody Response 4 Weeks After a 3-dose Primary Series to 4 Weeks After a Booster Dose	20	Change From Baseline in Imaging Specific Airways Resistance: siRaw Measured at FRC and TLC Scan Conditions, Presented in Scan Trimmed Scan Types, Measured in 5 Lobes and 5 Regions at Screening, Day 12 and Day 28	60
C08	Respiratory Tract Diseases	Comparison of Serotype-specific Geometric Mean Concentration of IgG Antibody Response 4 Weeks After a Booster Dose to One Year After a Booster Dose	20	PFS in Subgroups That Were Defined by Germline PDGFRB Polymorphisms	48
C09	Otorhinolaryngologic Diseases	Change From Baseline in HDM-specific IgE Levels at Week 8	4	Change From Baseline in AM and PM Peak Expiratory Flow Rate (PEFR)	11
C09	Otorhinolaryngologic Diseases	Pain Intensity (PI) as Rated on a 6-point VRS by the Patient at 0.5, 1, 2 and 3 Hours After the First Lozenge	3	Change From Baseline in Forced Expiratory Volume in One Second (FEV1)	3
C09	Otorhinolaryngologic Diseases	the Effect of Squeezable Bottle and Syringe on Clinical Effectiveness in Sinusitis Children	2	Change From Baseline in Forced Vital Capacity (FVC)	3
C09	Otorhinolaryngologic Diseases	Percentage of Participants Who Discontinued Study Drug Due to an Adverse Event	2	Change From Baseline in Forced Expiratory Flow (FEF) Between 25% and 75% of the Vital Capacity (FEF25%-75%)	3
C09	Otorhinolaryngologic Diseases	Total Nasal Airflow on Day 8 Measured Using Active Anterior Rhinomanometry (AAR)	2	Change From Baseline 24-hour Urinary Free Cortisol Level	3
C10	Nervous System Diseases	CSF IL-6, sTREM2, HMGB1, Albumin, IgG	10	Geometric Mean Titers (GMTs) of Meningococcal Serogroups A, C, Y, and W Antibodies Following Vaccination With 3 Lots of MenACYW Conjugate Vaccine	12
C10	Nervous System Diseases	ITP: Number of Combined Unique Active (CUA) Lesions, New or Enlarging Time Constant 2 (T2) Lesions, and New or Persisting Time Constant 1 (T1) Gadolinium Enhanced (Gd+) Lesions Per Participant Per Scan	6	Percentage of Participants Achieving hSBA Vaccine Seroreponse for Meningococcal Serogroups A, C, Y and W Following Vaccination With 3 Lots of MenACYW Conjugate Vaccine	10
C10	Nervous System Diseases	Mean Blood Flow Velocity in Middle Cerebral Artery	5	Incidence of Non-delirium Complications After Surgery	8

C10	Nervous System Diseases	Total Cumulative Prednisone Dose	4	Change From Baseline in Peak Work (in Watts/kg) During Exercise Testing at Week 12 in Part 1	7
C10	Nervous System Diseases	Statistically Significant Change From Baseline to 12 Week Endpoint in Laboratory Measures - Chloride, High Density Lipoprotein, Sodium, and Triglycerides	4	Change From Baseline in Respiratory Function Tests to Characterize the Degree of Involvement of Respiratory Muscles.	4
C11	Eye Diseases	Distribution of Change in Visual Acuity (Letters) From Baseline to 1 Year	8	Percentage of Ranibizumab Re-injections Received Over 28 and 52 Weeks	10
C11	Eye Diseases	Ex Vivo Total Cholesterol Uptake at Day 30	8	Change From Baseline in GA Area in Complement Factor I (CFI) Positive and Negative Participants at Week 48	7
C11	Eye Diseases	Mean Number of Macular Laser Treatments From Baseline Through Months 24 and 36	4	Percentage of Participants Avoiding a Loss of $\geq 15$ , $\geq 10$ , or $\geq 5$ Letters From the Baseline BCVA in the Study Eye Averaged Over Weeks 40, 44, and 48	6
C11	Eye Diseases	Change From Baseline in Visual Acuity (VA): Double Masked Phase	4	Distribution of Change in Visual Acuity (Letters) From Baseline to 1 Year	4
C11	Eye Diseases	Change From Baseline in Central Retinal Thickness (CRT) at Week 52 as Assessed on Optical Coherence Tomography (OCT) - LOCF	4	Loss of Visual Acuity	3
C12	Urogenital Diseases	Secondary Efficacy Endpoints - Vaginal Mucosa Assessment (Vaginal Color)	12	Systolic, Diastolic and Mean Blood Pressure at Week 0, 3, 6, 12, and 16	12
C12	Urogenital Diseases	Change From Baseline to Day 90 in Serum Potassium	7	Blastocyst Quality, Intention-to-treat (ITT) Analysis Set	7
C12	Urogenital Diseases	Change From Baseline in eGFR	7	Creatinine Clearance (CrCl) in the Immediate Post-transplant Period	6
C12	Urogenital Diseases	Progression-Free Survival (PFS)	5	Pharmacokinetics: AUC(0- $\infty$ ) of Tadalafil	6
C12	Urogenital Diseases	Percent Change From Baseline to 12 Week Endpoint in High-Density Lipoprotein Cholesterol (HDL-C) and Low-Density Lipoprotein Cholesterol (LDL-C)	4	Change From Baseline to 12 Week Endpoint in Total Cholesterol and Triglycerides	5
C13	Female Urogenital Diseases and Pregnancy Complications	Secondary Efficacy Endpoints - Vaginal Mucosa Assessment (Vaginal Color)	12	Systolic, Diastolic and Mean Blood Pressure at Week 0, 3, 6, 12, and 16	12
C13	Female Urogenital Diseases and Pregnancy Complications	Change From Baseline to Day 90 in Serum Potassium	7	Blastocyst Quality, Intention-to-treat (ITT) Analysis Set	7
C13	Female Urogenital Diseases and Pregnancy Complications	Change From Baseline in eGFR	7	Creatinine Clearance (CrCl) in the Immediate Post-transplant Period	6
C13	Female Urogenital Diseases and Pregnancy Complications	Change From AURORA 1 Baseline (i.e., Month 0) in Urine Protein to Creatinine Ratio (UPCR)	4	Progression Free Survival (PFS)	5
C13	Female Urogenital Diseases and Pregnancy Complications	Change From AURORA 1 Baseline (i.e., Month 0) in Urine Protein	4	Change From Baseline to Day 90 in eGFR	5
C14	Cardiovascular Diseases	Change From Baseline in Daytime (6am to 10 pm) Mean Systolic Blood Pressure Measured by Ambulatory Blood Pressure Monitoring.	27	Inflammatory Mediator Levels, Interleukin-1b, Interleukin 6 and Tumor Necrosis Factor (TNF) (pg/ml)	14
C14	Cardiovascular Diseases	Change From Baseline in Daytime (6am to 10 pm) Mean Diastolic Blood Pressure Measured by Ambulatory Blood Pressure Monitoring.	24	Change From Baseline in Mean Ambulatory Blood Pressure During the Final 2, 4, and 6 Hours of the Dosing Interval at Week 4	12
C14	Cardiovascular Diseases	Change From Baseline in the Trough (22-24-hr) Mean Diastolic Blood Pressure Measured by Ambulatory Blood Pressure Monitoring.	23	Change From Baseline in Plasma Concentration of Serum N-terminal Pro-Brain Natriuretic Peptide (NT-proBNP) at Months 5, 9, 13, 17, 21, 25, 29, 33, 37, 42, 48 and Final Visit	12
C14	Cardiovascular Diseases	Change From Baseline in the Trough (22-24-hr) Mean Systolic Blood Pressure Measured by Ambulatory Blood Pressure Monitoring.	22	Change From Baseline in Urine Albumin-to-Creatinine Ratio at Months 5, 9, 13, 17, 21, 25, 29, 33, 37, 42, 48 and Final Visit	12
C14	Cardiovascular Diseases	Change From Baseline in the Nighttime (12 am to 6 am) Mean Systolic Blood Pressure Measured by Ambulatory Blood Pressure Monitoring.	22	Change From Baseline in the Nighttime (12 am to 6 am) Mean Diastolic Blood Pressure Measured by Ambulatory Blood Pressure Monitoring.	11
C15	Hemic and Lymphatic Diseases	Mean Change in Ferritin and Hepcidin From Baseline to the End of the Primary Efficacy Period	10	Mean Change in Iron and Total Iron Binding Capacity (TIBC) From Baseline to the End of the Primary Efficacy Period	7
C15	Hemic and Lymphatic Diseases	Change From Baseline in Hematocrit	8	Mean Change in Hematocrit and Reticulocytes From Baseline to the End of the Primary Efficacy Period	6
C15	Hemic and Lymphatic Diseases	Summary of Subgroup Analyses for Kaplan-Meier Estimates for Time to Death From Any Cause	7	Change From Baseline in Reticulocyte Count	6
C15	Hemic and Lymphatic Diseases	Mean Change in Red Blood Cell (RBC) Count and Absolute Reticulocyte Count From Baseline to the End of the Primary Efficacy Period	7	Progression-free Survival (PFS)	5
C15	Hemic and Lymphatic Diseases	Mean Change in Hemoglobin (Hb) Levels From Pre-treatment to the End of the Primary Efficacy Period	6	Summary of Subgroup Analyses for Kaplan-Meier Estimates for Time to Death From Any Cause	5
C16	Congenital, Hereditary, and Neonatal Diseases and Abnormalities	Change From Baseline in Serum Uric Acid Levels at Day 1, Day 3, Day 7, Day 11, Day 14 and Follow-up	44	Change From Baseline in Plasma Levels of Hypoxanthine at Day 1, Day 7, Day 14, and at Follow-up	36
C16	Congenital, Hereditary, and Neonatal Diseases and Abnormalities	Change From Baseline in Plasma Levels of Xanthine at Day 1, Day 7, Day 14, and at Follow-up	22	Change From Baseline in Plasma Levels of Xanthine at Day 1, Day 7, Day 14, and at Follow-up	20
C16	Congenital, Hereditary, and Neonatal Diseases and Abnormalities	Percent Change From Baseline in Serum Urate Levels at Week 28.	13	Fraction of Inspired Oxygen (FIO2) (Percent) During the First 24 h and up to Day 7	11
C16	Congenital, Hereditary, and Neonatal Diseases and Abnormalities	Percent Change From Baseline in Serum Urate Levels at Final Visit	13	Oxygen Requirement and Ventilatory Support -- SpO2/FIO2 Ratio	10
C16	Congenital, Hereditary, and Neonatal Diseases and Abnormalities	Absolute Change in Percent Predicted Forced Expiratory Volume in 1 Second (ppFEV1)	10	Change From Baseline in Peak Work (in Watts/kg) During Exercise Testing at Week 12 in Part 1	7
C17	Skin and Connective Tissue Diseases	Change From Baseline in Erythrocyte Sedimentation Rate (ESR) at Weeks 2, 4, 8, 12, 16, 20, 24, 32, 40, 48, and 52	27	Change From Baseline in Predictive Biomarkers: Amyloid A, Chemokine (C-C Motif) Ligand 17, Chemokine (C-X-C Motif) Ligand 13, Interleukin 6, Macrophage-Derived Chemokine	26
C17	Skin and Connective Tissue Diseases	Change From Baseline in C-reactive Protein (CRP) at Weeks 2, 4, 8, 12, 16, 20, 24, 32, 40, 48, and 52	24	Change From Baseline in C-Reactive Protein (CRP) at Weeks 1, 2, 4, 8, 12 and 16	18
C17	Skin and Connective Tissue Diseases	Change From Baseline in Percentage Body Surface Area at Week 2, 4, 8 and 12	8	Change From Baseline in Flow Cytometry: 6 Colour TB Natural Killer (NK) Panel- CD16+CD56+, CD19, CD3, CD3+CD4+	16
C17	Skin and Connective Tissue Diseases	Percentage of Participants Achieving Investigator's Global Assessment (IGA) Response of Clear (0) or Almost Clear (1) and a Reduction of Greater Than or Equal to $\geq 2$ Points From Baseline at Weeks 12, 16, 28, 40, and 52: Double-blind Period	8	Change From Baseline in Predictive Biomarkers: Chitinase 3 Like 1, Matrix Metalloproteinase 3 (MMP-3)	14
C17	Skin and Connective Tissue Diseases	Percentage of Participants Achieving Patient Global Assessment (PtGA) Response of 'Clear (0)' or 'Almost Clear (1)' and Greater Than or Equal to 2 Points Improvement From Baseline at Weeks 12, 16, 28, 40 and 52: Double-blind Period	8	Change From Baseline in Flow Cytometry: CD16+ Monocyte Panel: CD14-HLA-DR+CD11cbr+CD123-, CD14br+CD16+, CD14br+CD16-, CD14lo+CD16br+	14

C18	Nutritional and Metabolic Diseases	Change From Baseline in Serum Uric Acid Levels at Day 1, Day 3, Day 7, Day 11, Day 14 and Follow-up	44	7-Point Self-Monitored Blood Glucose (SMBG) Profiles	53
C18	Nutritional and Metabolic Diseases	Change From Baseline in Fasting Plasma Glucose (FPG)	34	Change From Baseline in Fasting Lipid Profile (Triglycerides/Cholesterol)	39
C18	Nutritional and Metabolic Diseases	Change From Baseline in Carbon Monoxide Diffusion Capacity (DLco)	27	Change From Baseline in Plasma Levels of Hypoxanthine at Day 1, Day 7, Day 14, and at Follow-up	36
C18	Nutritional and Metabolic Diseases	Change From Baseline in Fasting Plasma Glucose	22	Change From Baseline in Carbon Monoxide Diffusion Capacity (DLco)	26
C18	Nutritional and Metabolic Diseases	Change From Baseline in Plasma Levels of Xanthine at Day 1, Day 7, Day 14, and at Follow-up	22	ADDENDUM: Insulin Dose	22
C19	Endocrine System Diseases	Change From Baseline in Fasting Plasma Glucose (FPG)	34	7-Point Self-Monitored Blood Glucose (SMBG) Profiles	53
C19	Endocrine System Diseases	Change From Baseline in Carbon Monoxide Diffusion Capacity (DLco)	27	Change From Baseline in Carbon Monoxide Diffusion Capacity (DLco)	26
C19	Endocrine System Diseases	Change From Baseline in Fasting Plasma Glucose	22	ADDENDUM: Insulin Dose	22
C19	Endocrine System Diseases	Percent Change From Baseline in ANGPTL3, TC, LDL-C, HDL-C, VLDL-C, Non-HDL-C, ApoB (ApoB-48, ApoB-100), ApoCIII, ApoAI, FFA, and Lp(a) at Primary Analysis Time Point	20	Change From Baseline in Fasting Plasma Glucose	21
C19	Endocrine System Diseases	Change From Baseline in Forced Expiratory Volume in One Second (FEV1)	19	INITIATION: 7-point Self-monitored Plasma Glucose (SMPG) Profiles and Postprandial Excursions	20
C20	Immune System Diseases	Change From Baseline in Fraction of Exhaled Nitric Oxide (FeNO) Over Time Following the Cessation of Repeat Dose Treatment With FF/VI	34	Change From Baseline in Peak Expiratory Flow (PEF) During Treatment and Following Cessation of Repeat Dose Treatment With FF/VI	30
C20	Immune System Diseases	Change From Baseline in Morning and Evening Peak Expiratory Flow Rate (PEFR) Over 26 and 52 Weeks of Treatment	28	Change From Baseline in Predictive Biomarkers: Amyloid A, Chemokine (C-C Motif) Ligand 17, Chemokine (C-X-C Motif) Ligand 13, Interleukin 6, Macrophage-Derived Chemokine	26
C20	Immune System Diseases	Change From Baseline in Erythrocyte Sedimentation Rate (ESR) at Weeks 2, 4, 8, 12, 16, 20, 24, 32, 40, 48, and 52	27	Change From Baseline in C-Reactive Protein (CRP) at Weeks 1, 2, 4, 8, 12 and 16	18
C20	Immune System Diseases	Change From Baseline in C-reactive Protein (CRP) at Weeks 2, 4, 8, 12, 16, 20, 24, 32, 40, 48, and 52	24	7-Point Self-Monitored Blood Glucose (SMBG) Profiles	18
C20	Immune System Diseases	Ratio to Baseline in Renal Biomarkers-Urine and Serum Beta-2 Microglobulin (B2M), Urine Albumin/Creatinine, Urine B2M/Urine Creatinine, Urine Phosphate, Urine Protein/Creatinine, Urine RBP 4 and Urine RBP 4/Urine Creatinine at Weeks 24, 48	22	Change From Baseline in Flow Cytometry: 6 Colour TB Natural Killer (NK) Panel- CD16+CD56+, CD19, CD3, CD3+CD4+	16
C23	Pathological Conditions, Signs and Symptoms	Cumulative Percentage of Participants With Treatment Failure	72	Cumulative Percentage of Participants With Treatment Failure	46
C23	Pathological Conditions, Signs and Symptoms	Pain Intensity Difference (PID)	47	Change From Baseline in Fasting Lipid Profile (Triglycerides/Cholesterol)	39
C23	Pathological Conditions, Signs and Symptoms	Time-weighted Sum of Pain Intensity Difference (SPID)	18	Pain Intensity Difference (PID)	20
C23	Pathological Conditions, Signs and Symptoms	Percentage of Participants With Treatment Failure	15	Non-opioid Rescue Medication - Ibuprofen	16
C23	Pathological Conditions, Signs and Symptoms	Time to Treatment Failure	12	Non-opioid Rescue Medication - Paracetamol	15
C24	Occupational Diseases	Total Sleep Time Measured by Actigraphy (Nights 6-7)	1	Latency to Persistent Sleep Over a 2-night Average Measured by Polysomnography (Nights 6-7)	3
C24	Occupational Diseases	Sleep Efficiency Measured by Actigraphy (Nights 6-7)	1	Latency to Persistent Sleep Over a 2-night Average Measured by Polysomnography (Nights 13-14)	3
C24	Occupational Diseases	Sleep Latency Measured by Actigraphy (Nights 6-7)	1	Total Sleep Time Over a 2-night Average Measured by Polysomnography (Nights 6-7)	3
C24	Occupational Diseases	Sleep Time Measured by Actigraphy (Nights 6-7)	1	Total Sleep Time Over a 2-night Average Measured by Polysomnography (Nights 13-14)	3
C24	Occupational Diseases	Sleep Time Measured by Actigraphy (Nights 13-14)	1	Sleep Efficiency Over a 2-night Average Measured by Polysomnography (0-3 Hours) (Nights 6-7)	3
C25	Chemically-Induced Disorders	Percent Medication Adherence at 3-month Follow-up Assessment	1	Percentage of Days Used Drugs or Alcohol	2
C25	Chemically-Induced Disorders	Clinician-assessed Depression Rating at 3 Month Follow-up Assessment	1	Length of Hospital Stay	2
C25	Chemically-Induced Disorders	CD4+ Lymphocyte Count at 12-month Follow-up Assessment.	1	Assessment of Bioequivalence of Prototype Mini Lozenges With Nicorette Mini Lozenge by Measuring Area Under the Plasma Concentration Versus Time Curve From Time Zero to Time t (AUC [0-t])	2
C25	Chemically-Induced Disorders	Change in Serum 3 $\alpha$ -androstenediol Glucuronide	1	Assessment of Bioequivalence of Prototype Mini Lozenges With Nicorette Mini Lozenge by Measuring Area Under the Plasma Concentration Versus Time Curve Calculated From Time Zero to Infinity (AUC [0-inf])	2
C25	Chemically-Induced Disorders	Time to First Rescue-free Laxation (Following the First Dose of Study Drug)	1	Percent Abstinent From Tobacco at Week 32 (7 Day Point Prevalence)	1
C26	Wounds and Injuries	Sum of Pain Intensity Difference at Rest and on Weight Bearing Over 6 Hours on Day 1 and Over 2 Hours on Day 3	3	Sum of Pain Intensity Difference at Rest and on Weight Bearing Over 6 Hours on Day 1 and Over 2 Hours on Day 3	9
C26	Wounds and Injuries	Number of Patients With Cutaneous Bacterial Load After Surgery	1	Sum of Pain Intensity Difference (SPID) on Weight Bearing Over 3 Days (SPID WB0-3)	3
C26	Wounds and Injuries	Degree of Filling of the Lesion by Repair Tissue at 12 Months Through MRI.	1	Sum of Ankle Pain Intensity Difference on Weight Bearing Over 24 Hours After Dose 1 (SPID WB24)	3
C26	Wounds and Injuries	The Difference in Total Bacterial Counts Measured in Colony Forming Units (CFU) as Determined by Quantitative PCR Analysis.	1	Sum of Pain Intensity Difference at Rest Over 24 Hours on Day 1 (SPID R24)	2
C26	Wounds and Injuries	Number of Inpatient Operating Room Debridements in Surgically Dehisced Wounds	1	Local Dynamic Stability (Knee During Turning With the Prosthesis on the Inside of the Turn)	1
F03	Mental Disorders	Secondary Efficacy Endpoints - Vaginal Mucosa Assessment (Vaginal Color)	12	Vital Signs: Systolic and Diastolic Blood Pressure Levels	12
F03	Mental Disorders	CSF IL-6, sTREM2, HMGB1, Albumin, IgG	10	Mean Change From Baseline (CFB) at Day 21 in Neurocognitive Function as Determined by Central Nervous System Vital Signs (CNS-VS) Test Battery	9
F03	Mental Disorders	Percent Change From Baseline to 12 Week Endpoint in High-Density Lipoprotein Cholesterol (HDL-C) and Low-Density Lipoprotein Cholesterol (LDL-C)	4	Mean Change From Baseline (CFB) at Day 84 in Neurocognitive Function as Determined by Central Nervous System Vital Signs (CNS-VS) Test Battery	9
F03	Mental Disorders	Change From Baseline to 12 Week Endpoint in Fasting Lipid Parameters Including Lipoprotein Subclasses	4	Incidence of Non-delirium Complications After Surgery	8
F03	Mental Disorders	Target Engagement Assays: Change From Baseline in Filamin A (FLNA) Linkages to alpha7 Nicotinic Acetylcholine Receptor (alpha7nAChR) and Toll-like Receptor 4 (TLR4) in Subject Lymphocytes	4	Median Baseline and Change From Baseline in Body Mass Index (BMI) During Phase 3	6