Outcome Entity Extraction Annotation Guideline

1. Introduction

The goal of this guideline is to aid in creating a manually annotated corpus that will serve as the gold standard in training a machine learning (ML) model for outcome entity extraction. The outcome entity comes from the PICO framework (Population, Intervention, Comparison, Outcome) which assists in medical evidence formulation. The corpus for annotation comes from the abstracts of published Randomized Controlled Trial (RCT) studies. As this annotation focuses on outcome entity extraction, there will only be one type of tag (for outcome), though a sentence may contain multiple outcome entities. A detailed definition of an outcome entity is provided in the section below (Section 2), while full examples are provided in Section 4. Section 3 provides annotation logistics.

Because outcomes are defined by the study design, annotators will have access to the entire abstract and ClinicalTrials.gov specified Brief Summary and Primary and Secondary Outcomes Measures while annotating. This is described further in Section 3. Example annotations of full abstracts and their justification will be provided in Section 4 for this reason. However, for simplicity, in Section 2 only sentences will be provided, and the reader is asked to assume that highlighted entities are defined as outcomes per the study design.

2. Recognition of Outcome entity

2.1 What is an outcome?

The outcome of a study is the answer to the question: "What metrics are used to test the study hypothesis?" It constitutes the *measurable* result of a study. Thus, outcomes should only be annotated when they are *a part of the study design*. Outcomes are usually the dependent variable in a study, whose value depends on the treatment/intervention. Outcomes for RCT studies can include indicators of disease or symptom progression, recovery, safety, drug efficacy, and more. Outcome entities highlighted in text must have meaning outside of the context of the abstract, in other words, they must be able to stand alone. P-values, confidence intervals, and actual outcome observation values should never be included in outcome entities, and outcome entities should be contiguous (except for the exception of coordination ellipsis described in Section 2.3).

2.2. Modifiers

Modifiers are descriptive and *provide directionality* to the outcome entity but *should not be included in the outcome entity*. The following provides an example of an outcome entity highlighted in yellow (as it will be for the remainder of this guideline) and its modifier highlighted in purple.

> "The high-risk group displayed higher survival probability than the low-risk group." (1)

Here, the outcome entity is 'survival probability' while its modifier is 'higher'. Note how 'higher' is not included in the outcome entity, but instead serves to provide directionality to it. 'Survival probability' is a quantitative, measurable result, while 'higher' serves as a descriptor.

The following provides another example.

"Mean change from baseline in mRNA levels was lower for the DTG arm as compared to placebo."[1] (2)

Here, the modifier is 'lower' while the outcome entity is 'mean change from baseline in mRNA levels'. It is important to note that this outcome entity is structured differently from the example above, which just consisted of two nouns. The outcome entity includes 'mean' and the phrase 'change from baseline in'. The inclusion of words and phrases like these are discussed in Section 2.7 Rules 12 and 8, respectively. Right now, this example serves to illustrate that outcome entities can take many forms so long as they are identified by the study design. So here, if the study design specifies measuring the change from baseline in mRNA levels it is highlighted, where 'mean' adds specificity.

The following provides an additional example of the variety of structure in outcome entities.

"The duration of a high viral load was shorter with REGEN-COV than with placebo."[2] (3)

Here, clearly the 'duration of a high viral load' is an entity that a study could aim to measure. Modifiers regarding duration are typically 'long', 'slow', 'short', etc. which can be seen here with the modifier 'shorter'.

Note that words like increase, decrease, reduction, etc. are typically used as modifiers (and thus not included in the outcome entity) (Section 4 Example 2). However, there are instances where these words should be included in the outcome as they are essential to the outcome concept. See Section 4 Example 4 for an example of where 'improvement' should be included in the outcome.

2.3 Text Boundary

It is important to define the text boundary for an outcome entity. Note that as stated above, there are no restraints on the number of words that can be included in an outcome entity, or the overall structure of an outcome entity. However, conjunctions (and, or, as, but, etc.) *cannot be in an outcome entity*. An example is provided below.

"The efficacy and safety of nintedanib in Japanese patients were consistent with the overall INBUILD population."[3] (4)

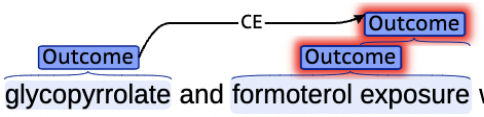
Here, efficacy and safety are considered two separate outcome entities, and are not grouped together. There is a conjunction 'and' that separates them.

There are times when 'or' is not used as a conjunction but more so as part of a mathematical symbol. Consider these cases as an exception to the above rule. An example is provided below:

"Grade 3 or worse adverse events were more frequent in the control group." (5)

Here, 'worse' has no meaning without its relationship with 'grade 3'. In a sense then, the 'or' acts to signify that they are looking for adverse events ≥ grade 3. Thus, in these specific cases 'or' may be included in the outcome entity.

Coordination ellipsis (CE) in text necessitates a specific way to mark the text boundary of an outcome entity. An example of this is as follows, which was taken from the Brat online platform.



[4](6)

Here, ‘glycopyrrolate exposure' and ‘formoterol exposure' are the two outcomes in this sentence. To use the annotation tool with this type of annotation, one should highlight the contiguous entity ‘formoterol exposure’ as one, and then annotate ‘glycopyrrolate’ with an arrow to the overlapping annotation of just ‘exposure’. The arrow is at ‘exposure’ because it comes last in the order of ‘glycopyrrolate exposure’. Note that cases of coordination ellipsis are the only time non-contiguous outcomes should be annotated and the only time that outcome annotations are allowed to overlap.

2.4 Other PICO Elements

It is important to contextualize the outcome entity annotation process. There are additional PICO elements, which, for practical annotation purposes, are the P (population) and I (intervention) elements (I and C (comparison) are typically grouped together in annotations). These elements *should not have any overlap* with the outcome entities. In other words, do not highlight any P or I elements as part of an outcome entity. The following provides an example of a sentence with the population element highlighted in blue and the intervention elements highlighted in green to make explicit the separation of these elements.

“A1C in individuals with impaired glucose tolerance was found to be lower with inhaled insulin vs. placebo.” (7)

Here, ‘A1C’ is the outcome entity, ‘individuals with impaired glucose tolerance’ is the population entity, and ‘inhaled insulin’ and ‘placebo’ are the two interventions.

2.5 Lists

When the term introducing the list is a clinical concept, and the introducing term and all elements of the list are a part of the study design, then the introducing term and its subsequent elements should all be highlighted as outcome entities. This is simply an example of redundant annotations which occur as per Section 2.7 Rule 1. An example is provided below.

“The counter-regulatory hormones glucagon, epinephrine, norepinephrine, growth hormone and cortisol were appropriately suppressed.”[5] (8)

Here, ‘counter-regulatory hormones’ is a complete clinical concept, which can be broken down into the elements ‘glucagon’, ‘epinephrine, ‘norepinephrine’, ‘growth hormone’ and ‘cortisol’, all of which are outcome metrics of the study. Thus, ‘counter regulatory hormones’ and all of the elements in the list are highlighted as outcome entities.

Sometimes the object introducing the list is too vague to act as an outcome entity on its own (recall the requirement that an outcome entity be able to "stand alone") and the elements are a part of the study design. The following provides an example.

"Patients experienced the primary endpoints of nausea, flu-like symptoms, and headaches." (9)

Here, 'primary endpoints' should not be marked as an outcome entity since that is both vague and not measurable. However, 'primary endpoints' serves to indicate to annotators that 'nausea', 'flu-like symptoms', and 'headaches' are all outcome metrics defined by the study design. Thus, each of 'nausea', 'flu-like symptoms', and 'headaches' are marked individually.

Additionally, sometimes in lists where the introducing term is an outcome entity, elements of the list were not specified as a part of the study design. This typically happens with patient-reported adverse events. When the object introducing the list is a clinical concept, and the elements of the list were not defined as part of the study design, the elements of the list should not be highlighted as outcome entities. An example is provided below.

"Patients experienced adverse events such as nausea, flu-like symptoms, and headaches." (10)

Here, 'adverse events' is a known, measurable outcome of many drug trials and thus is marked as the outcome entity. However, the elements of the list, namely, 'nausea', 'flu-like symptoms', and 'headaches' should not be highlighted as outcome entities as they are patient-reported and thus variable and unknown at the time of study design. They do not constitute as metrics used to test the study hypothesis.

Another example of outcome entity recognition in sentences with lists is provided below.

"A majority of the participants were white, female, and above the age of 30." (11)

Here, it is important to note that baseline characteristics of a study (that are not outcomes of the study) act as covariates or classification variables. Thus, none of the entities mentioned in the above sentence are outcomes as defined by the study design.

Similarly, sometimes multivariate analysis is used to find association between outcome entities and covariates. In these cases, only the outcome entities should be highlighted. An example of this is provided below.

"In a multivariate analysis, high plasma ribavirin concentration, low baseline hemoglobin, HCV genotype 1b, and IL28B genotype CC were associated with higher SVR12."[6] (12)

Here, the covariates are 'plasma ribavirin concentration, 'baseline hemoglobin', 'HCV genotype 1b', and 'IL28B genotype CC', which are all associated with the outcome entity 'SVR12', but should not themselves be highlighted as outcome entities.

2.6 Temporal Component

Outcome entities will occasionally have a temporal component that is essential to the meaning of the outcome. There are certain indicators to use to determine if the temporal component adjacent

to an outcome is essential to that outcome and should be highlighted. The indicators are as follows.

1. The temporal component is *explicitly stated by the study design.* This means that the temporal component is either, i) stated as part of the defined outcome in sections prior to the results and conclusions sections in the abstract or ii) stated as part of the defined outcome in the ClinicalTrials.gov file Primary or Secondary Outcomes Measures. Example 1 in Section 4 provides an example of this where '**adjusted mean change in HIV-1 RNA at day 8**'[1] is the outcome.

2. The temporal component is used as an adjective preceding the rest of the outcome entity. Typically, adjectival use of temporal components indicates that the inclusion of the temporal component adds clinical significance and meaning to the outcome. An example of this is as follows.

> "Changes from baseline with dapagliflozin 10 mg by day 7 were -2.29 mmol/L for **24-h daily average blood glucose**."[7] (13)

Here, '24-h' is the temporal component being used adjectivally preceding the rest of the outcome, namely, 'daily average blood glucose'. Thus, it is included as part of the outcome entity.

2.7 Additional Rules

1. Outcomes should be highlighted each time they are mentioned as an outcome
2. Outcomes beyond the scope of the current study should not be annotated as outcome entities
3. Articles preceding outcome entities should never be included in outcome entities
    a. Example: The **blood glucose level** increased.
4. Always highlight the outcome entities: **tolerability**, **efficacy**, and **safety**
    a. This includes all derivations of these words
        i. Example: **well-tolerated**, **tolerated**, **tolerability**
        ii. Example: **efficacious**
    b. Do not include any words that come after or before
        i. Example: **safety** profile
        ii. Example: **efficacy** parameter
        iii. Example: short-term **tolerability** variables
5. Always highlight the outcome entity: **adverse event(s)**
    a. This includes any subsets
        i. Example: **treatment-emergent adverse events**
        ii. Example: **grade 3/4 adverse events**
    b. Do not include anything else that comes after/before
        i. Example: **treatment-emergent adverse event** rates
        ii. Example: rates of **mild gastrointestinal adverse events**

6. Never highlight feasibility, effectiveness, superiority, or noninferiority (or their derivations) as stand-alone outcomes. They should not be treated the same as those mentioned in rule 4.
7. Do not highlight discontinuations or deaths as outcomes unless *explicitly stated by study design*
8. Only highlight 'changes in', 'differences in', 'improvement in', etc. if *explicitly stated by study design*
    a. Only highlight 'from baseline' if 'changes in', 'differences in', 'improvement in', or a similar entity is also included in the outcome
        i. Example: "At month 12, there was no significant difference in the change in serum Cr level from baseline."[8]
        ii. Example: "Once-daily 5% MTF resulted in differences in TAHC from baseline by 23.9 ± 2.1 hairs/cm at week 24."[9]
9. Example: Numeric values can be included in outcome entities if part of study design
    a. Example: recovery time to glucose level ≥3.9 mmol/l[5]
10. Parentheses (and the words/numbers that they encapsulate) adjacent to an outcome entity that are a part of the outcome entity (and not its observed value, CI, etc.) can be highlighted as a part of that outcome entity
    a. Example: recovery time to glucose level ≥3.9 mmol/l (≥70 mg/dl)[5]
11. An abbreviation of an outcome entity in parentheses immediately after the outcome entity should be highlighted separately as an outcome entity (redundancy)
    a. Example: adverse events (AEs)
12. Always try to include as much specificity in the outcome entity annotation as possible
    a. If specificity is provided by coordinate adjectives, commas can be included in the outcome entity
        i. Example: muscle invasive, p53-like bladder cancer progression
13. Phrases like 'need for', 'measures of', 'risk of', 'incidence of', 'occurrence of' generally shouldn't be highlighted as part of the outcome entity
14. Highlight outcomes at baseline if mentioned
    a. Example: mean baseline HbA1c
    b. Example: pain score at baseline

3. Logistics

The text corpus will be uploaded to the online platform called Brat. For abstracts with clear section labels, the user should annotate the results and conclusion section. If there are no clear labels, the user should use their best judgment to annotate the results and conclusion sections. The ClinicalTrials.gov fields Brief Summary, Primary Outcome Measures and Secondary Outcome Measures will be provided for each RCT abstract in a .txt file. Example .txt files are shown in Section 4.

The protocol for referencing ClinicalTrials.gov .txt files is as follows. Annotators should annotate outcome entities as identified by the content of the abstract. In cases when the annotator is unclear if a term is an outcome entity, the annotator should consult the ClinicalTrials.gov .txt

file. This protocol decreases the burden of cross-referencing on annotators and serves as a small safeguard against a potential incompleteness of outcomes listed on ClinicalTrials.gov. The inclusion of the reference ClinicalTrials.gov study design should make the annotation process clearer when there is uncertainty.

4. Full Abstracts

Note that as per Section 3, only the results and conclusions sections will be annotated. Additionally note that as per the protocol in Section 3, outcomes will be inferred from the context of the abstract unless otherwise stated.

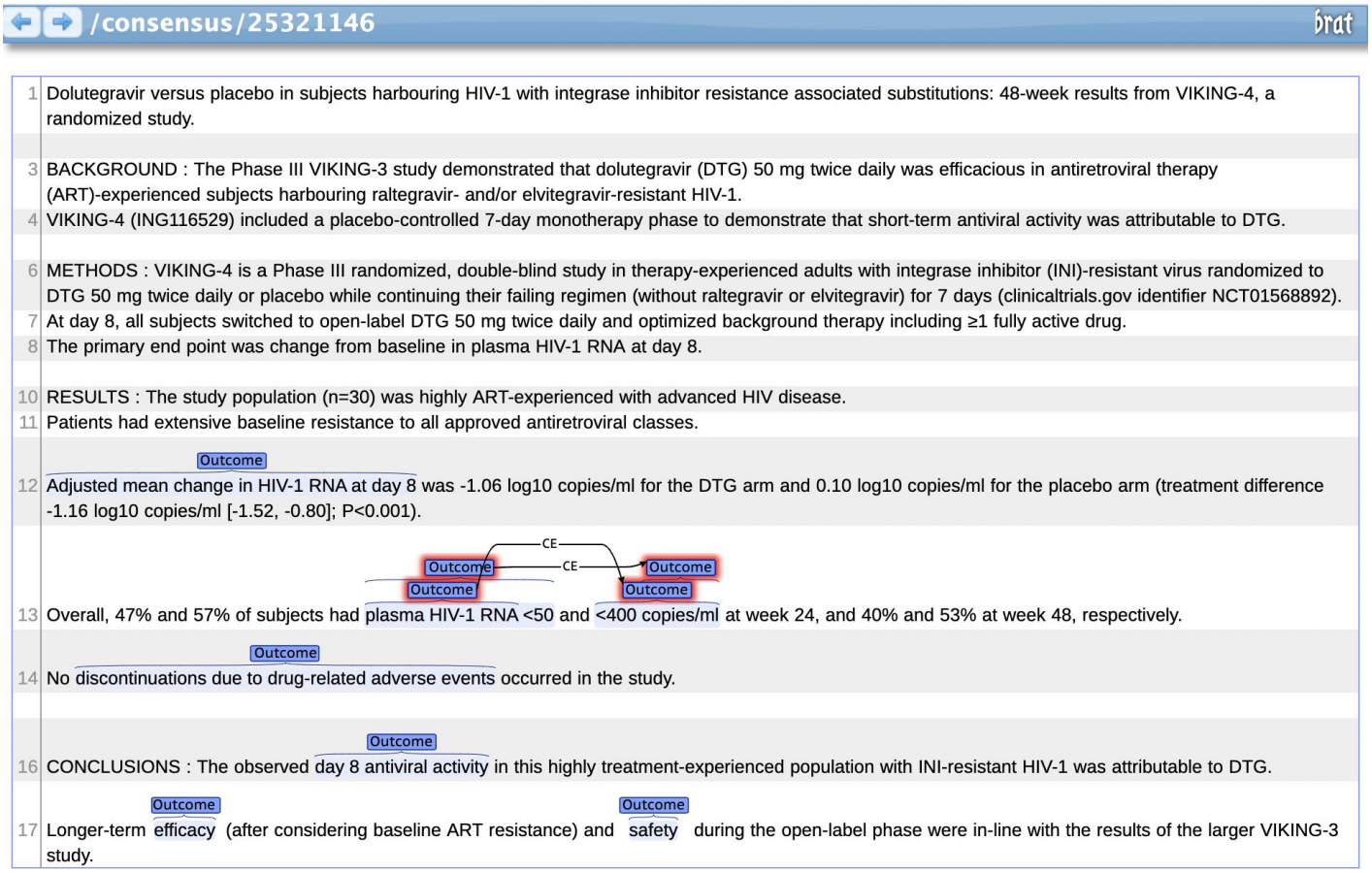
Example 1. PMID: 25321146[1]

ClinicalTrials.gov file:

25321146.txt

```
PMID:25321146 has 1 corresponding trials.


NCTID:NCT01568892

Brief Summary:Study ING116529 is a multicenter, randomized, study with an initial 7 day placebo- controlled, functional monotherapy
phase to quantify the antiviral activity attributable to dolutegravir (DTG) in HIV-1 infected, ART-experienced adults who are
experiencing virological failure on an Integrase inhibitor containing regimen (current RAL or ELV failures), with evidence of
genotypic resistance to RAL or ELV at study entry. Thirty subjects will be randomized (1:1) to receive either DTG 50mg BID (Arm A)
or Placebo (Arm B) with the current failing regimen for 7 days (RAL or ELV should be discontinued prior to dosing with DTG). At Day
8, subjects from both arms will enter an open label phase and receive open label DTG 50mg BID with an optimized background regimen
containing at least one fully active drug.

Primary Outcome Measures:
1. Mean Change From Baseline in Plasma Human Immunodeficiency Virus Type 1 (HIV-1) Ribonucleic Acid (RNA) at Day 8

Secondary Outcome Measures:
1. Absolute Values in Plasma HIV-1 RNA Over Time
2. Mean Change From Baseline in Plasma HIV-1 RNA Over Time
3. Number of Participants With Plasma HIV-1 RNA <50 c/mL Over Time
4. Number of Participants With Plasma HIV-1 RNA <400 c/mL Over Time
5. Absolute Values in Cluster of Differentiation 4+ (CD4+) Cell Counts Over Time
6. Median Change From Baseline in CD4+ Cell Counts Over Time
7. Absolute Values in Cluster of Differentiation 8+ (CD8+) Cell Counts Over Time
8. Median Change From Baseline in CD8+ Cell Counts Over Time
9. Number of Participants With the Indicated Type of HIV-1 Disease Progression (Acquired Immunodeficiency Syndrome [AIDS] or Death
[DT])
10. Number of Participants With Any Adverse Event (Serious and Non-serious) of the Indicated Grade
11. Number of Participants With the Maximum Post-Baseline-emergent Clinical Chemistry Toxicities of the Indicated Grade
12. Number of Participants With the Maximum Post-Baseline-emergent Hematology Toxicities of the Indicated Grade
13. AUC(0-tau) of DTG
14. Cmax of DTG
15. Plasma DTG Pre-dose Concentration (C0) at Day 8, Day 28, and Week 24; and Average DTG C0 (C0 Avg) at Week 24
16. Number of Participants With the Indicated Treatment-emergent Integrase (IN) Mutations Detected at the Time of Defined Virologic
Failure (PDVF), as a Measure of Genotypic Resistance
17. Number of Participants With the Indicated Fold Increase in Fold Change (FC) in the 50% Inhibitory Concentration Relative to
Wild-type Virus for DTG (i.e. PDVF FC/Baseline FC Ratio) at the Time of PDVF, as a Measure of Phenotypic Resistance
18. Number of Participants Who Discontinued Study Treatment Due to AEs
19. Number of Participants With Clinically Significant Electrocardiogram (ECG) Findings
20. Change From Baseline in Systolic Blood Pressure (SBP) and Diastolic Blood Pressure (DBP)
21. Change From Baseline in Heart Rate
22. Change From Baseline in Albumin Level
23. Change From Baseline in Alkaline Phosphatase (ALP), Alanine Aminotransferase (ALT), Aspartate Aminotransferase (AST) and
Creatine Kinase
24. Change From Baseline in Total Bilirubin (T. Bil) and Creatinine Levels
25. Change From Baseline in Cholesterol, Chloride, Carbon Dioxide (CO2)/Bicarbonate (HCO3), Glucose, High Density Lipoprotein
Cholesterol, Potassium, Low Density Lipoprotein (LDL) Cholesterol, Sodium, Phosphorus, Triglycerides and Urea/Blood Urea Nitrogen
26. Change From Baseline in Creatinine Clearance
27. Change From Baseline in Lipase Levels
28. Change From Baseline in Basophils, Eosinophils, Lymphocytes, Monocytes, Total Neutrophils, Platelet and White Blood Cell (WBC)
Count
29. Change From Baseline in Hemoglobin Level
30. Change From Baseline in Hematocrit Level
31. Change From Baseline in Mean Corpuscle Volume
32. Change From Baseline in Red Blood Cell Count
```

Abstract (in Brat):



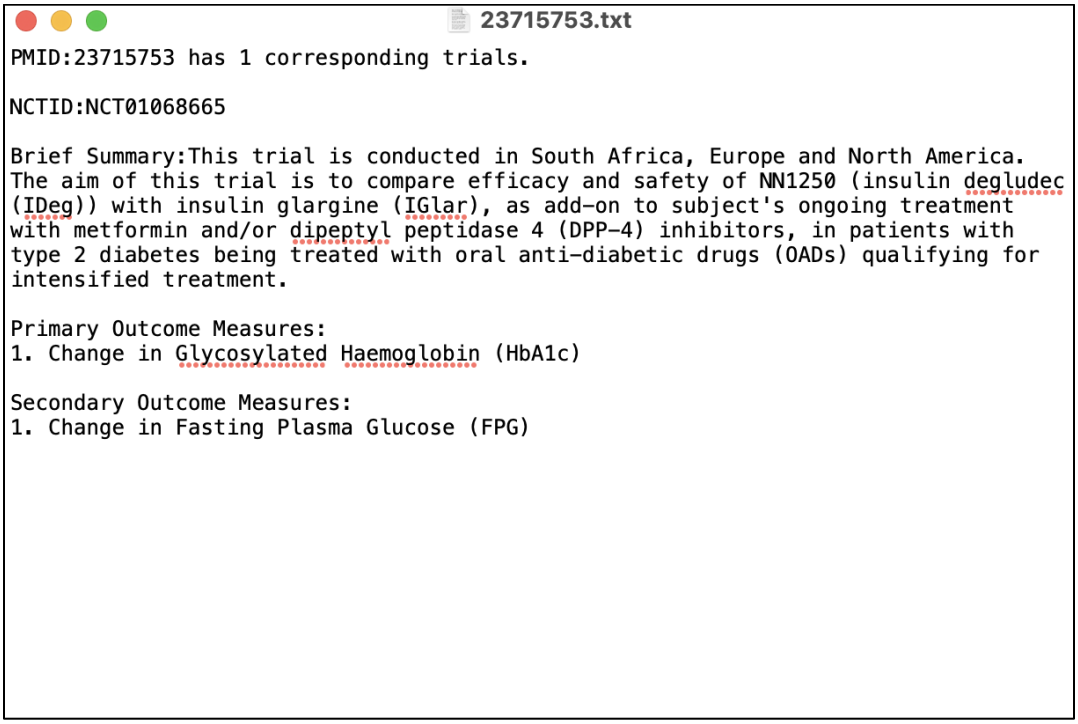
Justification:

Sentences 10 and 11 in the Results section provide baseline characteristics about the study population and thus nothing in these sentences is highlighted as an outcome entity. Sentence 12 consists of the outcome entity ‘adjusted mean change in HIV-1 RNA at day 8’. From the context of the abstract (especially line 7), the primary outcome involves measuring the change in HIV-1 RNA where ‘day 8’ and ‘change in’ and are defined as a part of the outcome entity in line 7, satisfying Section 2.6 and Section 2.7 Rule 8, respectively. ‘Adjusted’ serves to increase the specificity of the outcome entity and thus is included as per Section 2.7 Rule 12. If there was any lack of clarity, this outcome entity can be further confirmed by Primary Outcome Measure 1 in the ClinicalTrials.gov file. This process exemplifies the protocol of first inferring outcomes from the semantics of the abstract and if there is lack of clarity, subsequently checking the ClinicalTrials.gov file. ‘Treatment difference’ in line 12 is not highlighted as an outcome entity as it is not stand alone (i.e. treatment difference *in what*). In line 13, the two outcome entities ‘plasma HIV-1 RNA <50 copies/ml’ and ‘plasma HIV-1 RNA <400 copies/ml’ are highlighted. Weeks 24 and 48 are not *explicitly stated by the study design*, thus they are not essential components of the outcome and are not included. The inclusion of the values (copies/ml) is permitted per Section 2.7 Rule 9. The question of whether to include ‘discontinuations due to’ is addressed by Section 2.7 Rule 7 which requires ‘discontinuations due to’ to be *explicitly stated*

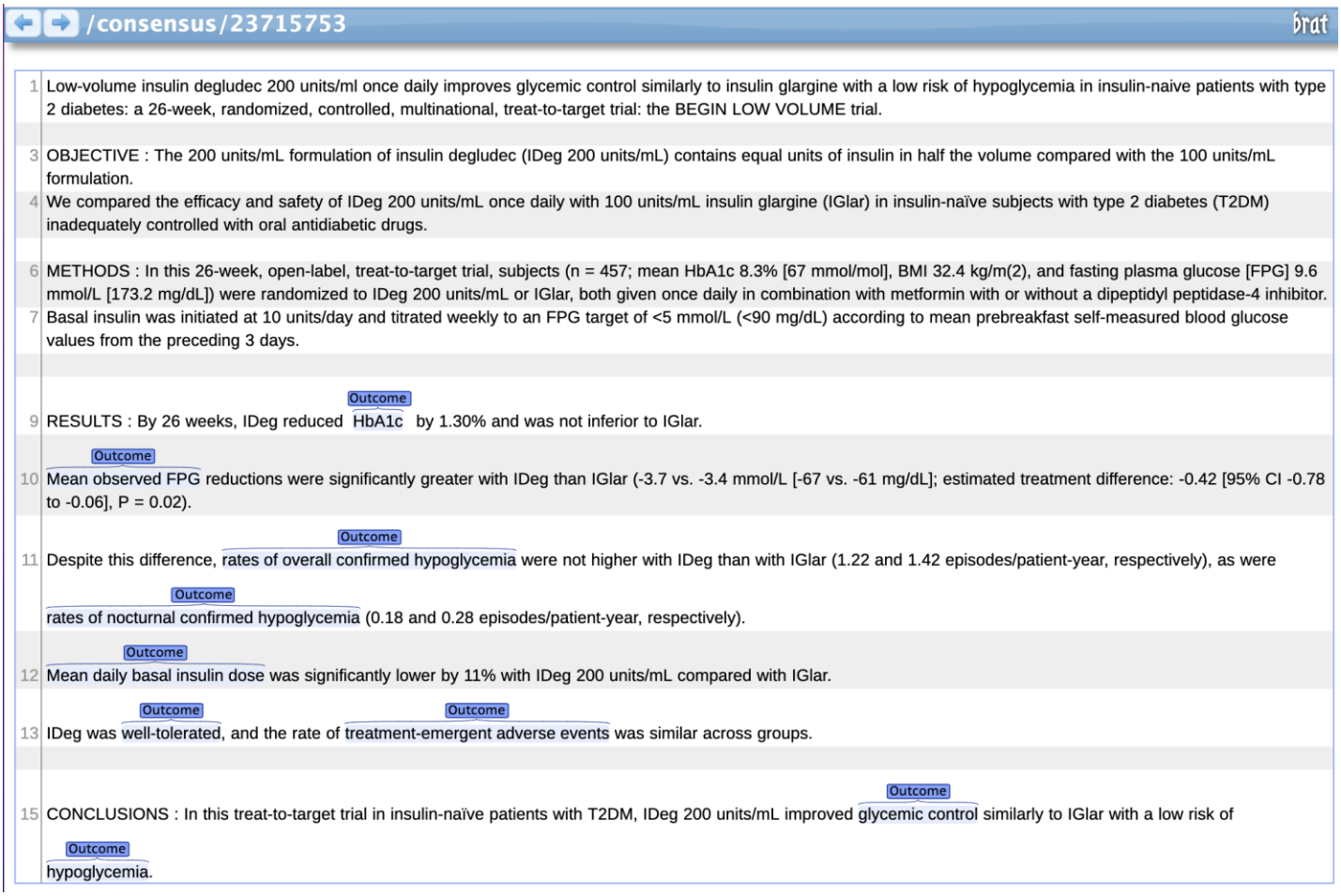
by the study design. While it is not mentioned in the background or methods sections, ‘number of participants who discontinued study treatment due to AEs’ is mentioned in Secondary Outcome Measure 18 in the ClinicalTrials.gov file. Thus, ‘discontinued due to’ is included in the outcome entity. In line 16, ‘day 8 antiviral activity’ is highlighted as an outcome entity as it is a more general term for the outcome highlighted in line 12. Lastly, ‘efficacy’ and ‘safety’ are highlighted in line 17 per Section 2.7 Rule 4.

Example 2. PMID: 23715753[10]

ClinicalTrials.gov file:



Abstract (in Brat):



Justification:

In line 9, ‘HbA1c’ is highlighted as an outcome entity. The word ‘reduced’ in front of ‘HbA1c’ acts as a modifier and thus is not included in the outcome entity. In line 10, ‘mean observed FPG’ is highlighted as the outcome entity. Here, ‘reductions’ is treated as a modifier and thus is not included in the outcome entity (it does not appear that they are interested in measuring ‘reductions’ specifically, just ‘change in’). The adjectives ‘mean observed’ are included in the outcome entity to add additional specificity as per Section 2.7 Rule 12. In line 11, ‘rates of overall confirmed hypoglycemia’ and ‘rates of nocturnal confirmed hypoglycemia’ are highlighted as outcome entities. ‘Rates of’ adds specificity to both entities and is included per Section 2.7 Rule 12. ‘Mean daily basal insulin dose’ is highlighted as an outcome in sentence 12, where ‘mean daily basal’ adds specificity (Section 2.7 Rule 12). In line 13, ‘well-tolerated’ is highlighted per Section 2.7 Rule 4, and ‘treatment-emergent adverse events’ is highlighted per Section 2.7 Rule 5 where ‘treatment-emergent' is included as it subsets ‘adverse events’ and ‘rate of’ is not included. In line 15, ‘glycemic control’ and ‘hypoglycemia’ are highlighted as they act as general, concluding terms to refer to the more specific outcome entities mentioned in the results section. Clearly, the ClinicalTrials.gov file corresponding to this abstract is nowhere near as comprehensive as in the previous example. However, this example shows how inferences can be made from the abstract itself.

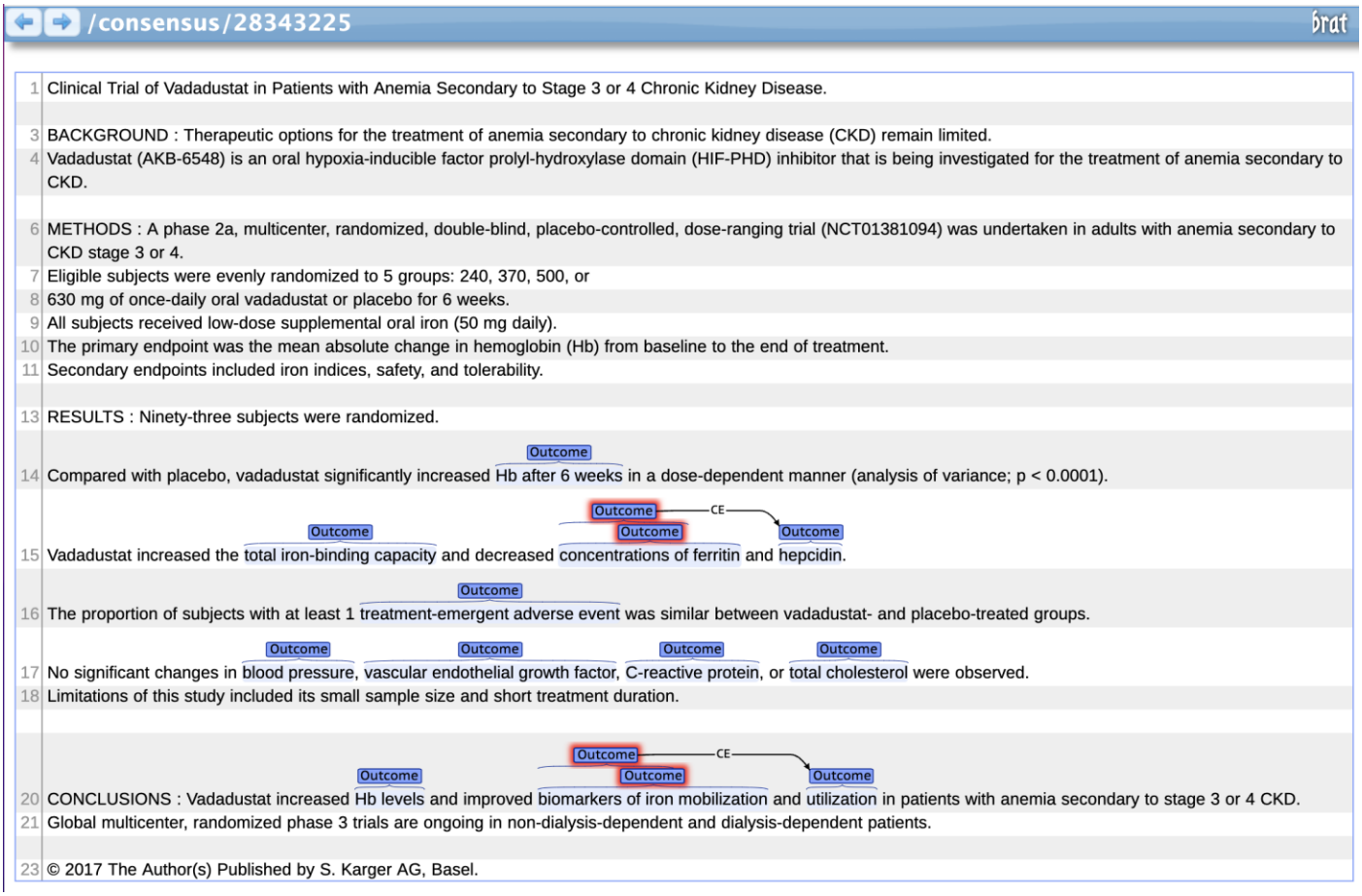
Example 3. PMID: 28343225[11]

ClinicalTrials.gov file:

28343225.txt

```
PMID:28343225 has 1 corresponding trials.

NCTID:NCT01381094

Brief Summary:The purpose of this study is to evaluate the dose response (efficacy), pharmacodynamic response, pharmacokinetics, safety,
and tolerability of orally administered AKB-6548 in pre-dialysis participants with anemia with repeat dosing for 42 days.

Primary Outcome Measures:
1. Absolute Change From Baseline in Hemoglobin (Hgb) to End of Treatment (Week 6)

Secondary Outcome Measures:
1. Change From Baseline in Hgb at Week 1, Week 2, Week 4, Week 6, and Follow-up Visit (up to Week 8)
2. Change From Baseline in Hematocrit (HCT) at Week 1, Week 2, Week 4, Week 6, and Follow-up Visit (up to Week 8)
3. Change From Baseline in Red Blood Cell (RBC) Count at Week 1, Week 2, Week 4, Week 6, and Follow-up Visit (up to Week 8)
4. Change From Baseline in Absolute Reticulocyte Count at Week 1, Week 2, Week 4, Week 6, and Follow-up Visit (up to Week 8)
5. Change From Baseline in Reticulocyte Hgb Content at Week 6
6. Maximum Change From Baseline in Hgb
7. Maximum Change From Baseline in HCT
8. Maximum Change From Baseline in RBC Count
9. Maximum Change in Absolute Reticulocyte Count From Baseline
10. Number of Participants With Absolute Change From Baseline in Hgb ≥ 0.4, 0.6, 0.8, and 1.0 g/dL at the End of Dosing Period
11. Number of Participants With Change From Baseline in Hgb ≥5.0, 7.5, and 10.0% by the End of Dosing Period
12. Number of Participants With Change From Baseline in HCT ≥5.0, 7.5, and 10.0% by the End of Dosing Period
13. Number of Participants With Change From Baseline in RBC Count ≥5.0, 7.5, and 10.0% by the End of Dosing Period
14. Number of Participants With Change From Baseline in Reticulocyte Count ≥6000, 12000, and 18000 Cells/uL by the End of Dosing Period
15. Change From Baseline in Total Iron at Week 2, Week 4, Week 6, and Follow-up Visit (up to Week 8)
16. Change From Baseline in Unsaturated Iron Binding Capacity at Week 2, Week 4, Week 6, and Follow-up Visit (up to Week 8)
17. Change From Baseline in Iron Saturation at Week 2, Week 4, Week 6, and Follow-up Visit (up to Week 8)
18. Change From Baseline in Total Iron Binding Capacity (TIBC) at Week 2, Week 4, Week 6, and Follow-up Visit (up to Week 8)
19. Change From Baseline in Ferritin at Week 2, Week 4, Week 6, and Follow-up Visit (up to Week 8)
20. Change From Baseline in Erythropoietin at Week 2, Week 6, and Follow-up Visit (up to Week 8)
21. Change From Baseline in Hepcidin at Week 6
22. Mean Plasma Vadadustat Concentrations on Week 2 and Week 4
23. Mean Plasma Vadadustat Acyl-Glucuronide Concentrations on Week 2 and Week 4
24. Number of Participants Treatment-emergent Adverse Events (TEAEs) and Serious Adverse Events (SAEs)
25. Number of Participants With Clinically Significant Changes From Baseline in Vital Signs Parameter
26. Number of Participants With Clinically Significant Abnormal 12-Electrocardiogram (ECG) Findings
27. Mean Change From Baseline in PR Interval, QT Interval, QRS Interval, and QT Corrected (QTc) Interval
28. Number of Participants With Clinically Significant Changes From Baseline in Laboratory Parameter Values
```

Abstract (in Brat):



Justification:

Sentence 13 specifies details of the trial procedure and thus nothing is highlighted as an outcome. In sentence 14, ‘Hb after 6 weeks’ is highlighted as the outcome entity. ‘After 6 weeks’ is included because of its explicit mention in Primary Outcome Measure 1 as well as Secondary Outcome Measure 1. In line 15, ‘total iron-binding capacity’ is highlighted. Additionally, ‘concentrations of ferritin’ and ‘concentrations of hepcidin’ are highlighted using the coordination ellipsis highlighting procedure described in Section 2.3. The outcome ‘treatment-emergent adverse event’ is highlighted in line 16 as per Section 2.7 Rule 5. In line 17, ‘blood pressure’, ‘vascular endothelial growth factor’, ‘C-reactive protein’ and ‘total cholesterol’ are highlighted as outcome entities. Note that ‘changes in’ is not included in the outcomes as it is not *explicitly stated by the study design*. Lastly, in sentence 20, ‘Hb levels’ is highlighted as it is a restatement of the outcome in line 14. Additionally, ‘biomarkers of iron mobilization’ and ‘biomarkers of utilization’ are highlighted via the coordination ellipsis highlighting procedure.

Example 4. PMID: 35286843[12]

ClinicalTrials.gov file:

35286843.txt

```
PMID:35286843 has 1 corresponding trials.

NCTID:NCT04317040

Brief Summary:This study evaluates the efficacy and safety of CD24Fc (MK-7110) in hospitalized adult
participants who are diagnosed with coronavirus disease 2019 (COVID-19) and receiving oxygen support.

The primary hypothesis of the study is clinical improvement in the experimental group versus the
control group.

Primary Outcome Measures:
1. Time to Improvement in Coronavirus Disease 2019 (COVID-19) Clinical Status
2. Number of Participants Who Experience an Adverse Event (AE)

Secondary Outcome Measures:
1. Percentage of Participants Who Died or Had Respiratory Failure (RF)
2. Time to Disease Progression in Clinical Status of COVID-19
3. Number of Participants Who Died Due to Any Cause
4. Rate of Clinical Relapse
5. Conversion Rate of COVID-19 Clinical Status
6. Time to Hospital Discharge
7. Duration of MV
8. Duration of Pressors
9. Duration of ECMO
10. Duration of High Flow Oxygen Therapy
11. Length of Hospital Stay
12. Change From Baseline in Absolute Lymphocyte Count
13. Change From Baseline in D-Dimer Concentration
```

Abstract (in Brat):

7 METHODS : We conducted a randomised, double-blind, placebo-controlled, phase 3 study at nine medical centres in the USA.
8 Hospitalised patients (age ≥18 years) with confirmed SARS-CoV-2 infection who were receiving oxygen support and standard of care were randomly assigned (1:1) by site-stratified block randomisation to receive a single intravenous infusion of CD24Fc 480 mg or placebo.
9 The study funder, investigators, and patients were masked to treatment group assignment.
10 The primary endpoint was time to clinical improvement over 28 days, defined as time that elapsed between a baseline National Institute of Allergy and Infectious Diseases ordinal scale score of 2-4 and reaching a score of 5 or higher or hospital discharge.
11 The prespecified primary interim analysis was done when 146 participants reached the time to clinical improvement endpoint.
12 Efficacy was assessed in the intention-to-treat population.
13 Safety was assessed in the as-treated population.
14 This study is registered with ClinicalTrials.gov, NCT04317040.

16 FINDINGS : Between April 24 and Sept 22, 2020, 243 hospitalised patients were assessed for eligibility and 234 were enrolled and randomly assigned to receive CD24Fc (n=116) or placebo (n=118).
17 The prespecified interim analysis was done when 146 participants reached the time to clinical improvement endpoint among 197 randomised participants.
18 In the interim analysis, the 28-day clinical improvement rate (Outcome) was 82% (81 of 99) for CD24Fc versus 66% (65 of 98) for placebo; median time to clinical improvement (Outcome) was 6·0 days (95% CI 5·0-8·0) in the CD24Fc group versus 10·0 days (7·0-15·0) in the placebo group (hazard ratio [HR] 1·61, 95% CI 1·16-2·23; log-rank p=0·0028, which crossed the prespecified efficacy boundary [α=0·0147]).
19 37 participants were randomly assigned after the interim analysis data cutoff date; among the 234 randomised participants, median time to clinical improvement (Outcome) was 6·0 days (95% CI 5·0-9·0) in the CD24Fc group versus 10·5 days (7·0-15·0) in the placebo group (HR 1·40, 95% CI 1·02-1·92; log-rank p=0·037).
20 The proportion of participants with disease progression (Outcome) within 28 days was 19% (22 of 116) in the CD24Fc group versus 31% (36 of 118) in the placebo group (HR 0·56, 95% CI 0·33-0·95; unadjusted p=0·031).
21 The incidences of adverse events (Outcome) and serious adverse events (Outcome) were similar in both groups.
22 No treatment-related adverse events (Outcome) were observed.
24 INTERPRETATION : CD24Fc is generally well tolerated and accelerates clinical improvement (Outcome) of hospitalised patients with COVID-19 who are receiving oxygen support.
25 These data suggest that targeting inflammation in response to tissue injuries might provide a therapeutic option for patients hospitalised with COVID-19.
27 FUNDING : Merck & Co, National Cancer Institute, OncoImmune.
29

Justification:

Sentences 16 and 17 describe the trial design. In sentence 18, ‘28-day clinical improvement rate’ and ‘median time to clinical improvement’ are highlighted as outcomes, as indicated by line 10. Note here that ‘28-day’ is an adjective to ‘clinical improvement rate’ and thus is included as per

Section 2.6. In line 19, ‘median time to clinical improvement’ is again highlighted as an outcome (redundancy). In line 20, ‘disease progression’ is highlighted as an outcome but ‘within 28 days’ is excluded as it is not *explicitly stated by the study design*. In line 21, ‘adverse events’ and ‘serious adverse events’ are highlighted as outcomes as per Section 2.7 Rule 5 (‘serious’ serves to subset ‘adverse events’). Similarly, ‘treatment-related adverse events’ is highlighted in line 22. Lastly, in line 24, ‘tolerated’ is highlighted as per Section 2.7. Rule 4 and ‘clinical improvement’ is highlighted because it is a redundant mention of an outcome (Section 2.7. Rule 1).

References:

1. Akil B, Blick G, Hagins DP, et al. Dolutegravir versus Placebo in Subjects Harbouring HIV-1 with Integrase Inhibitor Resistance Associated Substitutions: 48-Week Results from VIKING-4, a Randomized Study. *Antiviral Therapy*. 2015;20(3):343-348. doi:10.3851/IMP2878

2. O’Brien MP, Forleo-Neto E, Musser BJ, et al. Subcutaneous REGEN-COV Antibody Combination to Prevent Covid-19. *New England Journal of Medicine*. 2021;385(13):1184-1195. doi:10.1056/NEJMoa2109682

3. Inoue Y, Suda T, Kitamura H, et al. Efficacy and safety of nintedanib in Japanese patients with progressive fibrosing interstitial lung diseases: Subgroup analysis of the randomised, double-blind, placebo-controlled, phase 3 INBUILD trial. *Respiratory Medicine*. 2021;187. doi:10.1016/j.rmed.2021.106574

4. Reisner C, Fabbri LM, Kerwin EM, et al. A randomized, seven-day study to assess the efficacy and safety of a glycopyrrolate/formoterol fumarate fixed-dose combination metered dose inhaler using novel Co-Suspension^{TM} Delivery Technology in patients with moderate-to-very severe chronic obstructive pulmonary disease. *Respir Res*. 2017;18(1):8. doi:10.1186/s12931-016-0491-8

5. Hompesch M, Jones-Leone A, Carr MC, et al. Albiglutide does not impair the counter-regulatory hormone response to hypoglycaemia: a randomized, double-blind, placebo-controlled, stepped glucose clamp study in subjects with type 2 diabetes mellitus. *Diabetes, Obesity and Metabolism*. 2015;17(1):82-90. doi:10.1111/dom.12398

6. Asselah T, Zeuzem S, Soriano V, et al. ITPA Genotypes Predict Anemia but Do Not Affect Virological Response with Interferon-Free Faldaprevir, Deleobuvir, and Ribavirin for HCV Infection. *PLOS ONE*. 2015;10(12):e0144004. doi:10.1371/journal.pone.0144004

7. Henry RR, Rosenstock J, Edelman S, et al. Exploring the potential of the SGLT2 inhibitor dapagliflozin in type 1 diabetes: a randomized, double-blind, placebo-controlled pilot study. *Diabetes Care*. 2015;38(3):412-419. doi:10.2337/dc13-2955

8. Saag KG, Whelton A, Becker MA, MacDonald P, Hunt B, Gunawardhana L. Impact of Febuxostat on Renal Function in Gout Patients With Moderate-to-Severe Renal Impairment. *Arthritis Rheumatol*. 2016;68(8):2035-2043. doi:10.1002/art.39654

9. Efficacy and Safety of Once-Daily Minoxidil Foam 5% Versus Twice-Daily Minoxidil Solution 2% in Female Pattern Hair Loss: A Phase III, Randomized, Investigator-Blinded Study. JDDonline - Journal of Drugs in Dermatology. Accessed March 20, 2023. https://jddonline.com/articles/efficacy-and-safety-of-once-daily-minoxidil-foam-5-versus-twice-daily-minoxidil-solution-2-in-female-S1545961616P0883X/

10. Gough SCL, Bhargava A, Jain R, Mersebach H, Rasmussen S, Bergenstal RM. Low-volume insulin degludec 200 units/ml once daily improves glycemic control similarly to insulin glargine with a low risk of hypoglycemia in insulin-naive patients with type 2 diabetes: a 26-week, randomized, controlled, multinational, treat-to-target trial: the BEGIN LOW VOLUME trial. *Diabetes Care*. 2013;36(9):2536-2542. doi:10.2337/dc12-2329

11. Martin ER, Smith MT, Maroni BJ, Zuraw QC, deGoma EM. Clinical Trial of Vadadustat in Patients with Anemia Secondary to Stage 3 or 4 Chronic Kidney Disease. *Am J Nephrol*. 2017;45(5):380-388. doi:10.1159/000464476

12. Welker J, Pulido JD, Catanzaro AT, et al. Efficacy and safety of CD24Fc in hospitalised patients with COVID-19: a randomised, double-blind, placebo-controlled, phase 3 study. *Lancet Infect Dis*. 2022;22(5):611-621. doi:10.1016/S1473-3099(22)00058-5