Review of endpoint/outcome measures ontologies

Criteria used: coverage, structure, completeness, accuracy, fitness for purpose

General comments

- Trial protocols are riddled with acronyms that are sometimes hard to pick up with NER due to their ambiguous nature
- When capturing the acronyms, it is important to add the full label and preferably also a definition to avoid misunderstanding and ambiguous annotations
- Synonyms and definitions are beneficial for all classes in an ontology to enrich it and greatly enhance recall (annotation and annotation precision). Both ontologies would benefit from this.
- Use public ontologies where appropriate
    - The ontology would benefit from stable public ID:s, already available definitions and synonyms (no need to “reinvent the wheel”).
- Use public sources for definitions and synonyms such as Wikipedia
    - For HbA1c (https://en.wikipedia.org/wiki/Glycated_hemoglobin
- Safety endpoints such as “adverse events”, “SAEs” etc could be re-used in both ontologies as these tend to be the same or very similar regardless of indication and intervention.

GI_cancer ontology

Structure

- Nice hierarchy with branches for various outcome measure types

Completeness and accuracy

- Adding more synonyms would improve recall (including pluralisation, ADA vs ADAs, import of terms from public ontologies)
- Missing some commonly used terms in study protocols and study plans (in studies sampled from clinical trials.gov)
- Useful to look for official guidelines to augment the ontology with terms and synonyms (link below to FDA and RECIST guidelines)
- There are endpoint/outcome measures that are disease specific (cancer/tumour type specific) but there will be primary and secondary outcomes that are required for most solid tumor (interventional) clinical trials including GI cancer. These can be found in guidelines like RECIST (Response Evaluation Criteria in Solid Tumours) which is referred to in most trials and also FDA guidelines such as "Clinical trial Endpoints for oncology drug approval
- Synonyms and definitions have been added to many terms (including some synonyms and definitions imported from NCIT)
    - Would be useful to use more terms from public ontologies where possible (NCIT is a vocabulary/thesaurus specifically created to cover the cancer domain and contains a lot of useful terms)

Examples of "missing" terms:

- Complete Response | complete remission|CR
- Partial Response | partial remission
- Time to response | TTR
- Progressive Disease | PD
- Stable Disease
- Disease-free survival | DFS |r elapse-free survival | RFS
- Sum of diameters | SOD
- Duration of clinical benefit | DoCB
- Pathologic complete response | pCR

- Health-related quality of life | HRQoL
- Patient reported outcomes
- Adverse event of special interest | AE of special interest | AESI
- Treatment-emergent adverse events |TEAE | TEAEs (in ontology: Treatment-related adverse events which is not the same)
- Common Terminology Criteria for Adverse Events | CTCAE
- Immune related response criteria | immune response criteria (part of RECIST guildlines)
- In the ontology, SAE has the label Serious AE. This is more commonly called Severe adverse event, should be synonym if not class label

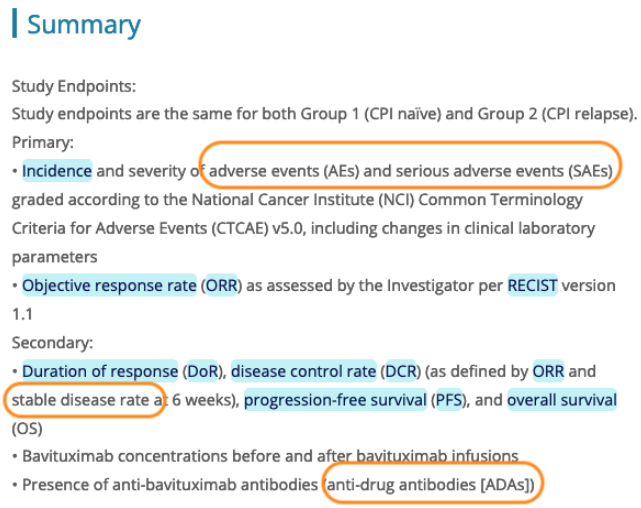
Examples of useful synonyms to add:
- Clinical benefit rate Duration of Clinical Benefit
- Overall survival OS
- Decrease of tumor size tumor shrinkage

Duplication in ontology:
- Dose limiting toxicity
- Dose-limiting toxicities

Fit for purpose
- With the addition of some more classes and synonyms/definitions, this would be useful for a GI cancer endpoint ontology and as a starting point for a solid tumor clinical trial endpoint/outcome measure ontology

Links to a few of the trials reviewed used for evaluation:
- https://clinicaltrials.gov/study/NCT00780494
- https://clinicaltrials.gov/study/NCT03675737
- https://clinicaltrials.gov/study/NCT04099641
- https://clinicaltrials.gov/study/NCT03472365
- https://clinicaltrials.gov/study/NCT02689284
  - Appenix 7 and 8 in the study protocol contains adapted RECIST 1.1 guidelines and Immune-Related RECIST guidelines https://cdn.clinicaltrials.gov/large-docs/84/NCT02689284/Prot_000.pdf
- https://www.fda.gov/media/71195/download Clinical trial Endpoints for oncology drug approval, FDA guidelines

Screenshot from TERMite UI and result annotations (blue) missed annotations (orange box)



Diabetes ontology

Structure

- Hierarchy looks good

Completeness and accuracy

Many of the most commonly used terms captured but annotation accuracy poor due to:

- Non-appropriate labels, missing synonyms and definitions
- Examples of not useful labels:
  - DiabeticNeuropathy, DiabeticFoot – won't match in text
  - AUROC, CGM, GAD-7 – no definition or full name
- To enrich the ontology, looking guidelines, papers, study protocol etc (as above for GI cancers) would help
- Also use public ontologies; for example GAD-7 is found with label, synonyms and definition in NCIT (http://purl.obolibrary.org/obo/NCIT_C103519 Generalized Anxiety Disorder - 7 Questionnaire)

Common endpoints in trials reviewed and not found in ontology:

- HbA1c is found in the ontology but the primary endpoint is most trials is to do with change in the level of HbA1c (glycosylated haemoglobin):
  - Reduction in HbA1c, Change in HbA1c, Change from baseline in HbA1c at 26 weeks as compared to placebo, Change from baseline to week 26 in glycosylated haemoglobin (HbA1c), glycemic control as measured by HbA1c etc
- Reduction in FPG, Change in FPG, fasting plasma glucose, fasting blood glucose
  - In ontology: "*FastingBloodGlucose",* "*Fasting plasma glucose level"* and no synonyms
- Reduced/increased body weight
- Anti-glucagon antibodies, anti-insulin antibodies (AIAs)
- Increase in PG (plasma glucose) change in PG (in ontology plasma glucose level and no synonyms)
- 8-point self-monitored plasma glucose (SMPG) self-measured plasma glucose (SMPG), 7-point SMBG
  - SMBG in ontology but no synonyms, definition
- time in range (TIR): TIR in ontology but no definition:
- Continuous glucose monitoring (CGM): CGM in ontology but no synonym or definition
- Clinically important hypoglycemic episodes
- Mixed meal tolerance test (MMTT): acronym very commonly used is missing
- Insulin Questionnaire Satisfaction Treatment (ITSQ)
- Residual beta cell function (RBCF)
- Change in C-peptide concentration/level
- Post prandial glucose (PPG), postprandial glucose (PPG) test: "*Post-prandialGlucose"* in ontology
- Injection site reaction: in GI cancer but not Diabetes ontology
- Number of hypoglycaemic episodes, hypoglycemia episodes
- Change in fasting and postprandial blood sugars
- Anti-drug Antibody, Anti-drug Antibodies (to add to both ontologies)
- Plasma glucose recovery
- Major Hypoglycemic Event Rate
- Hypoglycemic Event Rate
- Total cholesterol TC
- Triglycerides TG
- Maximal glucose infusion rate GIRmax
-

Duplicate in ontology

- FastingBloodGlucose

Fit for purpose:

- Would benefit for more manual curation to be useful

- Good starting point though with many relevant terms captured

Links to a few of the trials reviewed used for evaluation:
https://clinicaltrials.gov/study/NCT01377844
https://clinicaltrials.gov/study/NCT02691247
https://clinicaltrials.gov/study/NCT03170544
https://clinicaltrials.gov/study/NCT02465515

Screenshot from TERMite UI and result annotations (blue) missed annotations (orange box)