

Annotation Comments

Annotated lines in the ICSD version 3 (lines in the dataset ICSD3.tsv):

- 18835-21250
- 22416-22742
- 25168-25995
- 26277-26434
- 27743-28324
- 28875-29003
- 30922-31266
- 31413-31489
- 32394-32820
- 33090-33197
- 35077-35522
- 35844-35879
- 36851-37206
- 37311-37631
- 39394-40004
- 40319-40385
- 41074-41431
- 41680-41790
- 42580-42915
- 43186-43443

These lines consist of the introduction to the chapter called "Sleep Related Breathing Disorders", and the title, "Alternate Names", introduction, "Diagnostic Criteria"-section, "Essential Features"-section, and "Predisposing and Precipitating Factors"-section to each of the subchapters in "Obstructive Sleep Apnea Disorders" and "Central Sleep Apnea Syndromes". These subchapters are:

- "Obstructive Sleep Apnea, Adult"
- "Obstructive Sleep Apnea, Pediatric"
- "Central Sleep Apnea with Cheyne-Stokes Breathing"
- "Central Apnea Due to a Medical Disorder without Cheyne-Stokes Breathing"
- "Central Sleep Apnea Due to High Altitude Periodic Breathing"
- "Central Sleep Apnea Due to a Medication or Substance"
- "Primary Central Sleep Apnea"
- "Primary Central Sleep Apnea of Infancy"
- "Primary Central Sleep Apnea of Prematurity"
- "Treatment-Emergent Central Sleep Apnea".

The annotated sections were chosen by relevance for diagnosis. There was also a strict time limit, so there had to be limits on the number of sections to be annotated. The selection was done by a non-clinician without medical experience, but rather an informatics student.

Annotated ad hoc by informatics student Lise Chen with one round of annotation, and reviewed quickly one time by the Lise Chen again.

In the fourth column and after, there are comments, and typos are marked with a string containing "typo".

Comments on Labelling Rules

- conditions consisting of binding words, such as "with" or "due" are still labelled as one condition.
 - "Sleep apnea with Cheyne Stokes Breathing" annotated as [Sleep apnea with Cheyne Stokes Breathing]
 - Singular entities: "CSA-CSB", "central sleep apnea with cheyne-stokes breathing"
- Not including "the" unless a part of a proper name.
- Not include adjectives describing an entity unless part of a proper noun.
- "sleep related [x]" like "sleep related hypoventilation disorders" treated as one entity.
- words that contain typos or are missing small parts, but make sense, are labelled as the word it is understood as.
- "of" included when perceived as part of disease name, like the understanding of "sleep apnea of prematurity" as one entity.
- a condition followed by the words "syndrome", "disorder", "disorders" or similar is understood as including the trailing "syndrome", "disorder", "disorders" or similar.
- words binded by or, or different things where binding words are used, making two entities not easily splitted up, or an abbreviation is inserted in the middle of the entity, are regarded as one entity. Examples of entities where this applies (these are interpreted as ONE entity): "nocturnal restriction or congestion", "American Academy of Sleep Medicine (AASM) Manual for the Scoring of Sleep and Associated Events", "prolonged and intermittent obstructions that disrupt normal ventilation during sleep, normal sleep patterns, or both", "adult and pediatric hypoventilation disorders".
- the annotator does not have any medical competence, and the distinction between what was a certified medical condition or just a medical concept was unclear. As a rule of thumb, the rule "I have _" was used to determine if something was a condition or a concept. Example of use of this rule: "I have apnea" does not sound correct, but "I have sleep apnea" sounds better. Therefore "apnea" was considered a concept, while sleep apnea was considered a condition.

Comments on the Labels

- Concept became a sort of catch-all label, as all of the entities deemed important, but didn't belong to any of the other groups fell into this category. Apneas and other medical terms should be better classified than just as "concept"
- many important things were not annotated. The lack of certain labels were discovered too late.
 - Some useful labels to add would be:
 - cause (especially with CSA due to high-altitude descent)
 - measurement (arterial PCO2 of x, HAT+++)

- timestamps
- characteristics of conditions (CA, OA, instability of ventilatory control, "reduction or cessation of airflow", "abnormal increase in the arterial PCO2")
- requirement for condition (minimum x years / months of age)
- these are marked with a comment in the fourth and fifth column (after three tabs in tsv file). Sometimes the comments contain suggestions to what could be a label, other times they contain just the word "something" to highlight the token that should have had an entity label.
- should have included "objective findings" because of words that could be used to better label test terminology.
- very few things annotated with "outcome". This should perhaps not have been included as a label.