

Annotator instructions for Interesting Features annotation

Instructions:

The following is a list of features, found to be predictive in predicting future onset of a specific disease at least 1 year prior to the disease's diagnosis. The population for all diseases is an adult cohort from the UK Biobank, partially controlled for BMI, gender and age. Features include medical diagnoses, lifestyle factors, test results, demographics, and questionnaires (e.g., diet). We want to find **interesting features**.

Each feature is accompanied by:

- **Feature name**
- **AI model explanation** (optional to consider, as the model's reasoning is not always robust)
- **Direction of correlation** with the target disease (e.g., positively or negatively correlated)

We need your expert judgment on how **novel, plausible, useful, and overall interesting** each feature is.

What to Do

Your task is to evaluate how:

1. **Novel** (Is this association new or unexpected?)
2. **Plausible/Makes sense** (Does it make sense based on current knowledge?)
3. **Useful/Utility** (Would it have practical or clinical relevance?)
4. **Overall Interesting** (Considering its novelty, plausibility, and utility)

the feature appears. You will assign a score for each criterion using a **1–4 scale**:

1. **Strongly Disagree**
2. **Disagree**
3. **Agree**
4. **Strongly Agree**

(For instance, “Novelty: 4” would mean you *Strongly Agree* this feature is novel.)

You may also add comments to clarify your rating and overall opinion, in the “Comments” column.

For example, for the overall “Interesting” rating, 1: “Not interesting at all”, 4: “Really interesting, e.g. would like to research it further; or is a feature I would want to present as an example in a paper”.

- **Feel free to ignore or only lightly use** the AI model explanations (and literature citations) provided with each feature.

Example Annotations

Below are **illustrative scenarios** showing how you might apply these 4-point ratings. Note how the scale is applied to each criterion:

Example 1

- **Disease:** Lung Cancer
- **Feature:** "Smoking nicotine," positively correlated
 - **Novelty:** 1 (Strongly Disagree that it's novel; we already know this link well)
 - **Plausibility:** 4 (Strongly Agree it is plausible; decades of evidence support it)
 - **Utility:** 3 (Agree it is useful; it's actionable for prevention, but also well-known)
 - **Overall Interestingness:** 1 (Strongly Disagree; it's too obvious to be interesting)

Example 2

- **Disease:** Lung Cancer
- **Feature:** "Smoking nicotine," **negatively** correlated
 - **Novelty:** 4 (Strongly Agree that it's novel; it contradicts established knowledge)
 - **Plausibility:** 1 (Strongly Disagree it's plausible; no known mechanism to support this)
 - **Utility:** 1 (Strongly Disagree it's useful; even if data said 'protective,' the broader health implications make it unlikely to be applied)
 - **Overall Interestingness:** 4 (Strongly Agree; if truly robust, this is very intriguing and worth deeper research)

Example 3

- **Disease:** Lung Cancer
- **Feature:** "Smoking cannabis," positively correlated
 - **Novelty:** 2 (Disagree; somewhat documented, though less extensively than tobacco)
 - **Plausibility:** 3 (Agree; smoking route is associated with lung irritation, potential carcinogens)
 - **Utility:** 2 (Disagree; might not immediately change clinical practice without more data)
 - **Overall Interestingness:** 2 or 3 (Disagree or Agree; it's somewhat interesting but not a big leap from known risks)

Expanded guidelines:

Rating Scale Definitions

Each criterion should be rated on a scale of **1 (Strongly Disagree) to 4 (Strongly Agree)**. Below are some general guidelines for interpreting the scale in each category:

1. Novelty

- **1 (Strongly Disagree):** Not novel at all; this association is obvious or firmly established.
- **2 (Disagree):** Slightly novel; mildly surprising, but there is some prior knowledge or literature.
- **3 (Agree):** Moderately novel; not extensively documented, raises interesting questions.
- **4 (Strongly Agree):** Highly novel; very surprising or challenges current literature/knowledge.
 - Example: (70 years ago): Lithium curing bipolar mania.

2. Plausibility/makes sense

- **1 (Strongly Disagree):** Not plausible; conflicts with well-established evidence or lacks a clear mechanism.
 - E.g. You would not feel comfortable presenting it to your boss as a finding
- **2 (Disagree):** Low plausibility; rationale is weak or uncertain.
- **3 (Agree):** Reasonably plausible; aligns with known mechanisms or partial evidence.
- **4 (Strongly Agree):** Very plausible; strongly supported by known biology, social factors, or established theories.
 - E.g. You would feel comfortable mentioning it to your boss or peers.

3. Utility (Usefulness)

- **1 (Strongly Disagree):** Not useful; offers no clear practical benefit or application.
- **2 (Disagree):** Slightly useful; may have niche relevance but limited broader impact.
- **3 (Agree):** Moderately useful; could inform some research or clinical decisions.
- **4 (Strongly Agree):** Highly useful; likely to have real-world impact (e.g., guiding interventions, policy, or significant new research).
 - Example: Exposure to sunlight and exercise improves health and outcomes

4. Overall Interestingness

- **1 (Strongly Disagree):** Not interesting at all; trivial, already well-known, or not worth further inquiry.
- **2 (Disagree):** Somewhat interesting; minor curiosity but probably no significant follow-up.
- **3 (Agree):** Moderately interesting; has enough novelty/plausibility/utility to prompt some investigation.
- **4 (Strongly Agree):** Very interesting; stands out as a new insight or provocative idea you'd want to research or present. Sparks curiosity or future research questions.
 - Example: People with a gradual, non-symptomatic decline in haemoglobin levels are at high risk of colorectal cancer. (Source: Dan Ofer/CRI).

Definitions:

Novelty:

Is the association between the feature and target novel, surprising, or not well-documented in current knowledge or literature? Consider also the correlation if relevant. Does this feature provide new insights or contradict established understanding? Is it Novel?

Oppositely, is it well known, established, or trivially explainable by other known factors? (i.e. "boring").

A hypothesis that introduces fresh perspectives or challenges existing views would score high for novelty.

Plausible/makes sense:

"Does it make sense for the feature to be associated with disease risk/onset based on known mechanisms, biological pathways, socioeconomic factors or theories? Is there a plausible explanation (or mechanism) for this relationship that makes sense?"

Utility/Usefulness

Usefulness: The practical relevance and applicability of each feature. Would it be beneficial for researchers, practitioners, or the general public? A hypothesis that provides actionable insights or has the potential to drive meaningful change would be rated high on this criterion.

Interestingness:

An interesting feature, should among other things, be novel, somewhat plausible (not embarrassing nonsense, like "Eating sparking rocks prevents insanity" [i.e. lithium and bipolar depression ;)]), have utility, and be the basis of usefulness.

Evaluate how *interesting* this feature is to a researcher, biologist, clinician or doctor. Consider world knowledge, analysis, vibes, and the criteria of novelty and plausibility.