

Annotation guideline: Managing side effects

(English version)

In this research project, you will help us by annotating forum posts from the Facebook group GIST International. To protect the privacy of the users, you will only be able to access the data via an interface and usernames have been removed.

With the data that you and other patients label, we will be able to develop a piece of software that can automatically extract both the negative side effects that patients experience and the strategies that they use to manage these side effects. If we can automatically extract this information, this could improve the quality of life of patients on the long term because patients and clinical researchers alike would gain insight into effective strategies for managing side effects that currently negatively impact quality of life.

Imagine that a patient would want to know how to manage his muscle cramps. It would not be doable for this individual patient to look through the entire forum for the right information and advice. A computer however would be well suited for this chore as long as it is able to recognize both the side effects and the strategies to manage them.

Therefore, in these Facebook messages we are looking for two types of concepts (also called entities):

1. Adverse drug responses (ADR)
2. Strategies for managing ADR

You can find more information below. In the examples the highlighted words are the correct annotation.

General guidelines for annotation

- Please annotate a whole word or set of words. Do not annotate half of a word. Try not to miss the first or last letter of a word.
- If a forum post is (half) in a different language, you can skip it.

Adverse Drug Responses (ADR)

An adverse drug response is a negative side effect of medication that a patient is using. Someone is shared their own experience or that of someone they know of a negative side effect they experienced (or are still experiencing). It does not count as an ADR if someone gives general information about side effects. You can however assume that if someone is asking for advice for dealing with a side effect, that they also have this side effect. In that case, it is an ADR.

If the sign or symptom is related to the disease instead of with the medication, it is not an ADR. Side effects of surgery also do not count as ADR. If it is related to reducing the dosage

or quitting the drug altogether, it is also not an ADR (then it is a withdrawal symptom). You can see examples of all these cases below.

Summary: Please label as ADR if it is a negative side effect resulting from medication use and it is clear someone has experienced this ADR.

Positive examples:

- Example of an ADR:

“I am so **nauseous** from taking Gleevec”

- Example of an ADR in a question. Here you can assume that this person also has the side effect:

“Anyone have suggestions for **nausea**? I am on 400mg Gleevec”

Negative examples

- Example that is not an ADR because it is a symptom of the disease itself:

“My abdomen hurts. My doctor says it is the GIST”

- Example that is not an ADR because it is associated with the patient quitting the drug or reducing dosage:

“Finally quit Gleevec. Now having a hard time sleeping”

- Example that is not an ADR because it is not clear if someone had this experience themselves:

“Nausea is a common side effect of Gleevec”

- Example that is not an ADR because it is a positive effect of medication:

“I feel so much better and have less headaches not that I’m using Gleevec”

Rules for annotating ADRs (ADR)

- Do not include articles (= lidwoorden) like ‘a’ and ‘the’ :

“I have a **headache**”

- Do not include possessive pronouns (= bezittelijke voornaamwoorden) like ‘my’ or ‘your’:

“My stomach hurts “

“I have pain in my leg “

- Do not include personal pronouns (= persoonlijke voornaamwoorden) like ‘I’ or ‘he’:

“He hardly feels human anymore “

“I could not sleep”

- Only include emotion-related words like ‘feel’ when the side effects related to subjective feelings (so not in the 1st example but only in the 2nd):

“I feel dizzy “

“I do not feel human “

- If someone uses ‘and’ or ‘or’ to show different options or different symptoms, please annotate them separately:

“I have a headache and a stomach ache “

- Duplicates should be annotated:

“The anxiety was debilitating. I also had a headache but the anxiety was worse.”

- Exclude similes and metaphors:

“Very hard to take a deep breath like someone is squeezing my lungs. “

- Functional problems, like problems with social activities or daily functioning, do count as ADRs.

“I would just stay around and do nothing all day.”

- Subjective complaints also do count as ADRs.

“It decreased my anxiety, but I hardly feel human anymore.”

- Do not include words that show how bad or severe the ADR is

“ My doctor increased my dose and I felt severe dizziness.”

- Do not include words that show how long the ADR lasted.

“I could not sleep for a whole week”

- Do include words that indicate where in the body the ADR took place:

“I have **pain** in my **leg**”

Complex cases : using the labels ADR-overlap and ADR-part

- Sometimes an ADR is split into two parts in different parts of the sentence. Please then label the separate parts with **ADR-part** instead of ADR like in the example below:

The image displays two examples of the labeling interface for Adverse Drug Reactions (ADRs). Each example consists of a list of labels at the top and a sentence below with words highlighted and labeled.

Example 1 (Top):

- Labels: ADR (a), ADR-overlap (c), STRATneg-part (m), STRATneg (n), STRATneg-overlap (o), STRATpos (p), STRATpos-part (q), STRATpos-overlap (r), ADR-part (b).
- Sentence: "Putting ice on my **leg** reduces the **pain**".

Example 2 (Bottom):

- Labels: ADR (a), ADR-overlap (c), STRATneg-part (m), STRATneg (n), STRATneg-overlap (o), STRATpos (p), STRATpos-part (q), STRATpos-overlap (r), ADR-part (b).
- Sentence: "I feel such an **ache** in my **head**".

- Sometimes one word in the sentence is part of two ADRs. However, in the interface you can only give a word one label. Therefore, we have added a special label for these cases **ADR-overlap**. Please annotate the shared word with the label **ADR-overlap** and the other parts of the two (or more) ADRs with the label **ADR-part**

Here is two examples:

5/19

ADR a ADR-overlap c STRATneg-part m STRATneg n
STRATneg-overlap o STRATpos p STRATpos-part q STRATpos-overlap r
ADR-part b

I have pain ✖ in my arm ✖ and leg ✖

< Prev Next >

6/19

ADR a ADR-overlap c STRATneg-part m STRATneg n
STRATneg-overlap o STRATpos p STRATpos-part q STRATpos-overlap r
ADR-part b

I have leg ✖ and stomach ✖ cramps ✖

< Prev Next >

Finally: Sometimes it is hard to decide whether a word does or does not belong to the ADR. Please try to follow these guidelines as best you can and give your best interpretation.

If you need help or have any questions, you can contact me at XXX

Strategies for managing ADR (STRAT)

Patients often advise other patients about which strategies to use for reducing adverse drug responses. Most of the time, they recommend strategies but sometimes they also advise against certain strategies. Sometimes patients also simply discuss the strategies that they use without giving a specific advice.

Strategies can be a wide range of things such as changing your lifestyle (like running); eating certain foods (like pickle juice) or taking other drugs (like nausea medication). There can also be many other types of recommendations that are possible like putting soap under your sheets against muscle cramps.

Please annotate all strategies that patients recommend others to do or not to do or in other words that they are giving a positive or negative advice for. You can assume that if patients say they are using a certain strategy without saying that it does not work, that they are recommending it i.e. giving a positive advice. If a patient is saying that the doctor recommended a strategy, you can also assume this is a positive advice unless the patient explicitly says that is it not or that it did not work for them.

Examples:

“ **Drinking pickle juice** really helps with my nausea”

“ I tried **running** every day but it did not help at all “

“I would try **eating mangoes** “

To make a difference between strategies that DID work and those that DO NOT please use **STRAT-pos** (positive strategy) for the ones that work and **STRAT-neg** (negative strategy) for the ones that do not. So:

“ **Drinking pickle juice** really helps with my nausea”

“ I tried **running** every day but it did not help at all “

“I would try **eating mangoes** “

Just like for ADRs, also STRATegies can be split into two parts or there can be words that are part of 2 or more STRATegies. Just like with ADRs, please use the -part and -overlap version of the labels for these cases.

Some examples:

The image displays two screenshots of a software interface for annotating text. Each screenshot has a header bar with a progress indicator and a page number (6/21 and 7/21 respectively). Below the header is a blue bar containing several colored buttons with labels: ADR (yellow), ADR-overlap (orange), STRATneg-part (pink), STRATneg (dark red), STRATneg-overlap (red), STRATpos (green), STRATpos-part (light green), STRATpos-overlap (dark green), and ADR-part (orange). The main area of each screenshot shows a text input field with a sentence. In the top screenshot, the sentence is "Do not take calcium or magnesium supplements." and the words "calcium", "magnesium", and "supplements" are highlighted in red, indicating a negative strategy. In the bottom screenshot, the sentence is "I always take calcium and magnesium supplements." and the words "calcium", "magnesium", and "supplements" are highlighted in green, indicating a positive strategy. Navigation buttons labeled "< Prev" and "Next >" are located at the bottom of each screenshot.

6/21

ADR a ADR-overlap c STRATneg-part m STRATneg n

STRATneg-overlap o STRATpos p STRATpos-part q STRATpos-overlap r

ADR-part b

Do not take calcium or magnesium supplements.

< Prev Next >

7/21

ADR a ADR-overlap c STRATneg-part m STRATneg n

STRATneg-overlap o STRATpos p STRATpos-part q STRATpos-overlap r

ADR-part b

I always take calcium and magnesium supplements.

< Prev Next >

Rules for annotating Strategies (STRAT):

- The message does not always include the ADR. It can also be a response in the discussion thread on that ADR.
- Do not include articles (= lidwoorden) like 'a' and 'the' unless they are in the middle of a strategy :

"You should try eating a banana"

"A hot bath helps me when I am feeling tired"

- Do not include possessive pronouns (= bezittelijke voornaamwoorden) like 'my' or 'your':

"My **magnesium pills** from the Etos are the best remedy"

"Putting **ice on** my **leg** helps "

- Do not include personal pronouns (= persoonlijke voornaamwoorden) like 'I' or 'he':

"I go **running** to reduce the pain"

- If someone uses 'and' or 'or' to show different options or different strategies, annotate them separately:

"I recommend **eating more fruit** or **more vegetables**"

"I recommend **drinking less alcohol** and **going to bed on time**"

- It is also possible that a strategy has two parts. If someone uses 'and' to indicate this then please annotate as one strategy and include the 'and'.

"I recommend **taking one pill in the morning and one at night**"

"I recommend combining **coffee and chocolate** against nausea"

- Duplicates should be annotated:

"In the past I tried **magnesium pills** and **extra potassium**. Only the **magnesium** worked for me"

Example of combined ADR & STRAT

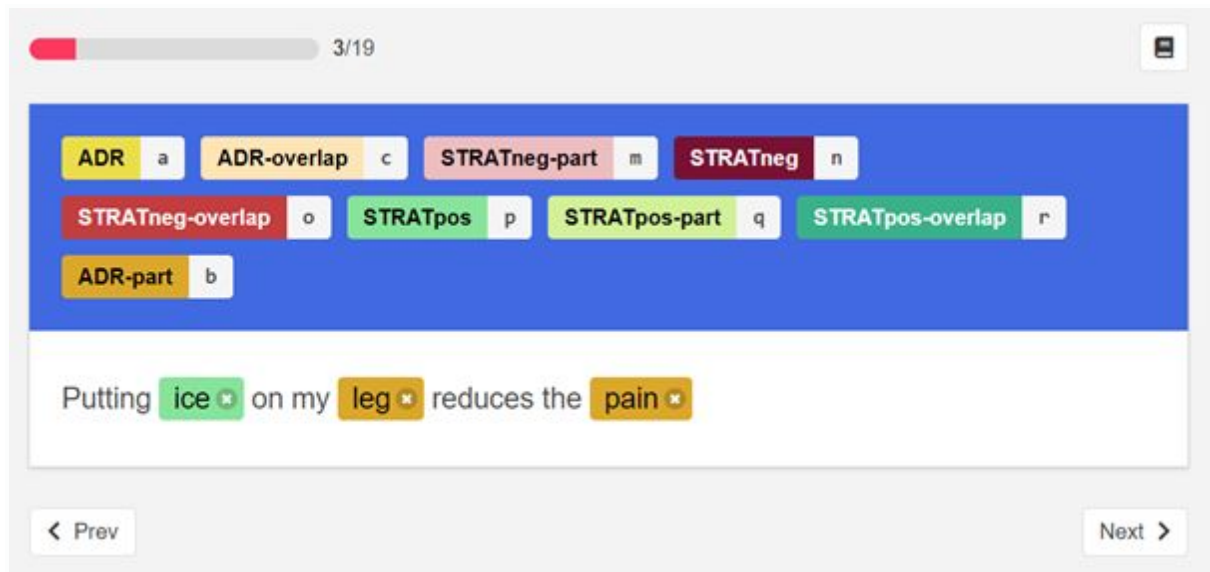
Sometimes you may encounter cases where a word could belong to both a ADR and a strategy. This is a difficult case.

"Putting ice on my **leg** reduces the **pain**"

And

"Putting **ice on my leg** reduces the pain "

In this case the word leg belongs to both an ADR and a strategy. Since you can only give one label to the word leg, there is no perfect solution. In this case, you will need to judge for yourself what the best possible combination is. For example I would label this as:



How do I use the interface to label posts?

Now you have read all the guidelines for annotating, I will explain how you can use the interface to start annotating the data.

First of all, please start Firefox or Chrome

The first step is to log in to Doccano: the online annotation system. Go to <http://textdata.nl/doccano>

If this does not work try: <http://textdata.nl:8000/login>

You now get a login screen. Use your first name (with a capital letter at the start and the rest in small letters) as username and "gistforum123" as password.

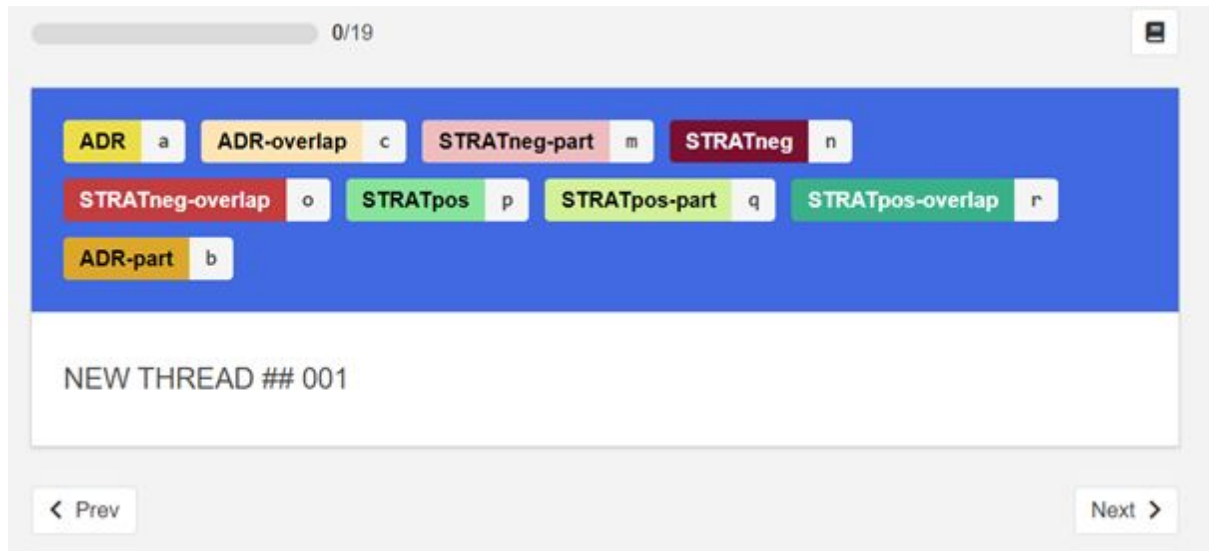
You are now logged in. You should see 2 projects. If this is not the case, you can contact me at XXX.

Please finish the top one first and then continue with the one with your name. Click on the project to start. You are now ready to begin labelling.

How to label ?

When you start the project, the first screen will say NEW THREAD ## some number.

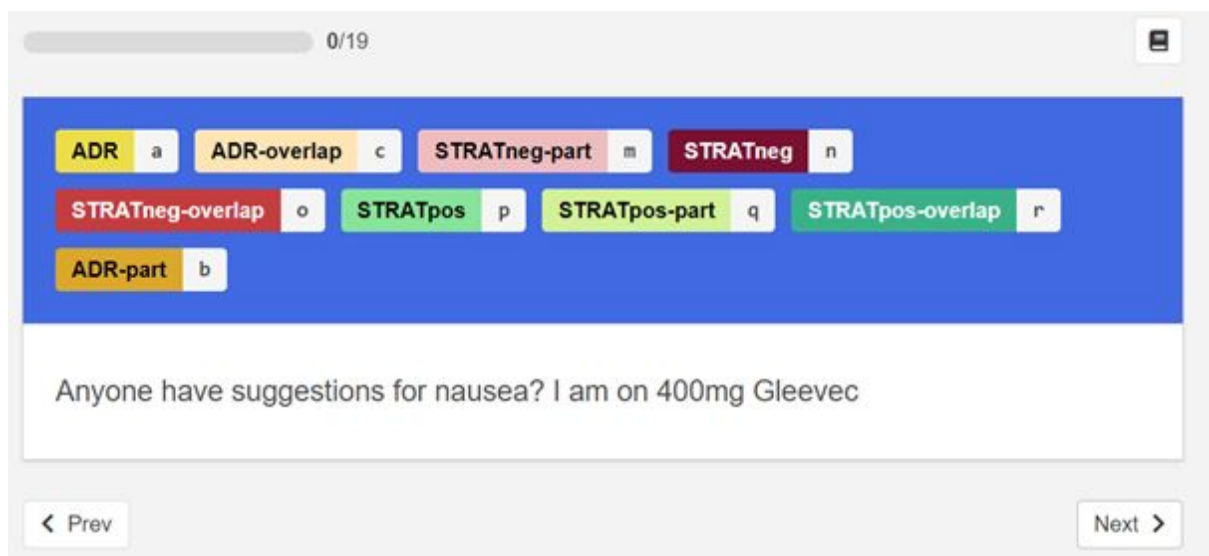
Like this one:



This indicates the start of a new forum discussion thread. This helps me keep track of which discussion you have worked on but it also helps you to know that a new discussion (which might have a new topic) has started.

Every time you see a screen which says NEW THREAD ## a number you can press the “Next” button in the bottom left to move to the first message of the discussion.

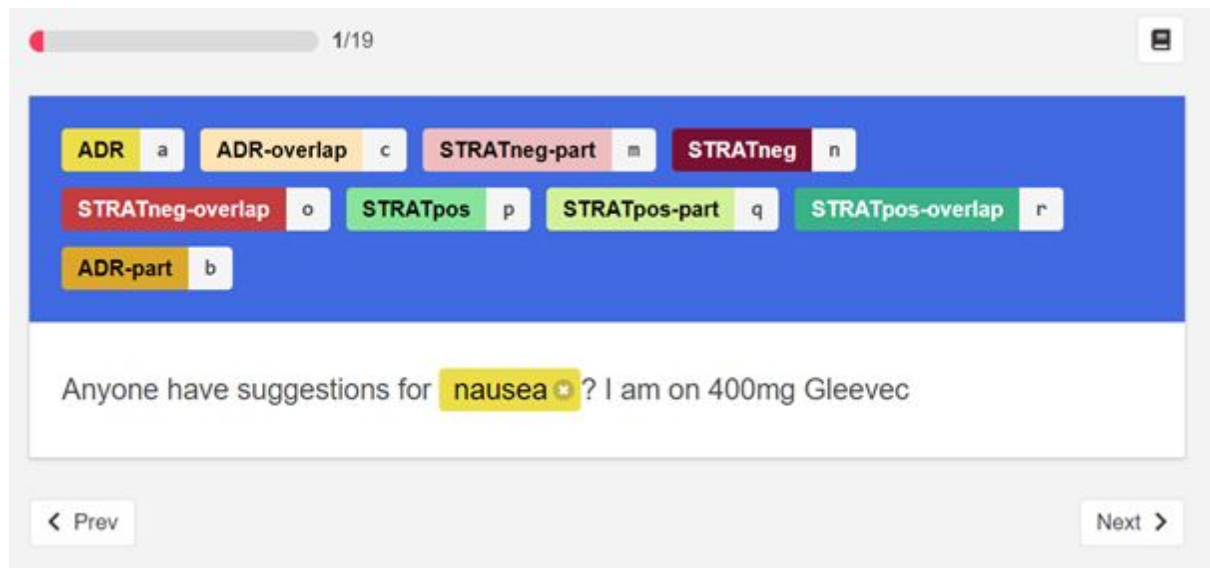
This is an example of what your screen will look like:



Please read the sentence carefully and if there is a negative side effect (ADR) or Strategy (STRAT) in there select the word (or words) and click on the correct concept type at the top.

A single word can also be double clicked to select it and instead of clicking on the concept type at the top you can also use the shortcuts. The shortcuts are the letters next to the concept type (like 'a' for ADR). This is a little faster.

After labelling it will look like this:



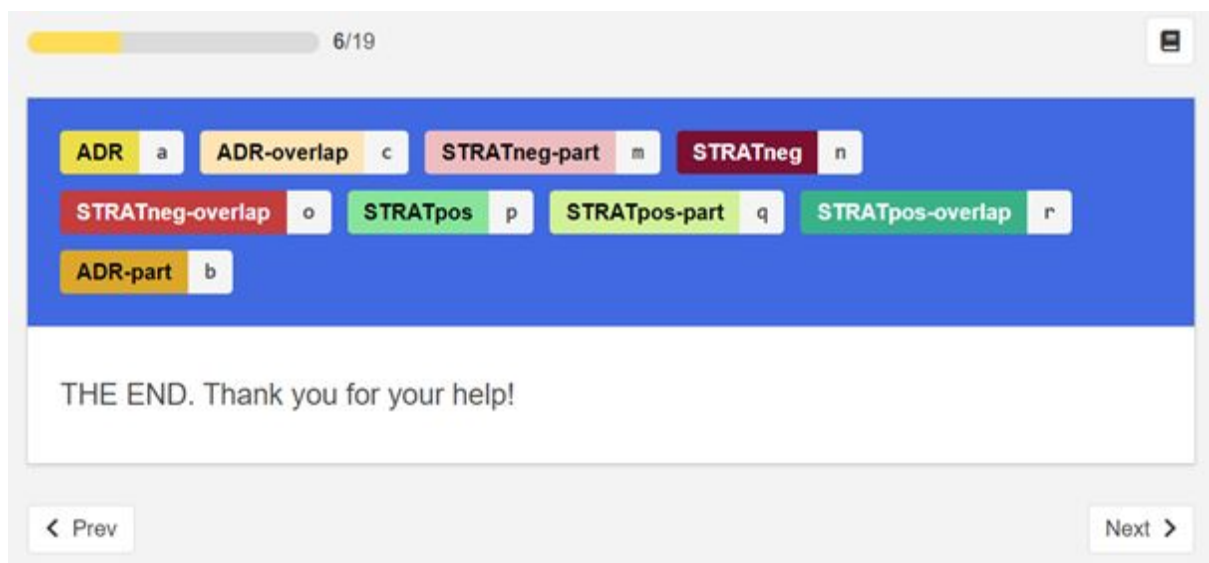
If you are done with a forum message, you can click on 'Next' to go to the next message of the discussion. It is also possible that a message does not contain any ADR and/or Strategy. In that case, go to the next message by clicking 'Next'. The next message might be:



This is a response to a question about strategies for managing nausea so in this context pickle juice is a STRAT-pos:



Please continue labelling until you reach the last sentence and get this screen:



Please note that the bar at the top does not give an accurate indication of where you are in the data because it only counts messages that you have given a label and not the ones that do not contain any ADR or strategy. So please ignore this bar.

IMPORTANT: If you want to stop or take a break, copy the link in the top of your browser. If you copy paste this in your browser later, then you can continue where you stopped last time. If you forget, do not worry. This is not a problem. It will just take some time because you will have to click through all the sentences that you have already labelled before you can continue.

IMPORTANT: If you are colorblind, have trouble with the colours of the labels or do not see the difference between two labels, let me know so I can change it for you.

Finally, thank you so much for wanting to help me out with my research! If you need any help or have questions, you can always email me at XXX. If we cannot solve it via email, we can also arrange a meeting near you home.