Proposal for enhanced qualifier classes

We here present some enhanced qualifier classes for Dutch clinical text. When processing clinical documents (e.g., doctor’s notes, discharge letters), detecting qualifiers (e.g., absent, historical, non-patient, increasing, decreasing, severe, light, etc.) follows the matching of concepts (e.g., diagnoses, symptoms, procedures). In this document we primarily use the term “qualifier”, although the terms “context”, “meta-annotation”, and “modifier” are also used to denote the same concept.

A consensus on potential qualifier classes, along with **clear** definitions of the qualifiers they encompass, is needed to develop accurate algorithms for detecting them. Despite some shared intuitive understanding of recognizing, for instance, a *negation* in a sentence, there are numerous cases where intuition simply falls short. In practice this impedes manual annotation of gold standards, used for training and evaluating algorithms. In turn the resulting annotations (with Kappa-Cohen as low as 0.5) lead to a difficult target for supervised machine learning models. A standardized classification of qualifiers, as proposed in draft form here, will hopefully advance both research and clinical implementation of NLP (Natural Language Processing) algorithms. We hope to receive as much feedback as possible before writing a definitive version and implementing it in [clinlp](https://github.com/umcu/clinlp).

For this first proposal we took the Context Algorithm (Harkema, 2009) as a starting point, both because this is an influential paper, and because a corresponding Dutch corpus is available (Afzal, 2014). There are already some trained models available that can for a large part be re-used. We will here describe three qualifier classes: Presence, Temporality, and Experiencer, including some definitions, issues to resolve, and illustrative examples. These classes can be further de-aggregated at a later stage. Other classes may follow later as well.

Qualifier classes are denoted by underlining, with the *qualifiers* (the mutually exclusive values a qualifier class can assume) italicized.

# Presence

|  |  |  |
| --- | --- | --- |
| ***absent/negated*** | ***uncertain*** | ***present/affirmed*** (default) |
| Concepts that are explicitly or implicitly described as absent in a patient | Whether the concept was absent or present is not asserted with very high certainty | Concepts that are explicitly or implicitly described as present in a patient |

Assessing whether some concept was present or absent is one of the most important parts of a clinician’s job. Whether something is present or occurred in the real world is knowable in principle, but in the clinical world, such assertions are rarely made with complete certainty. This is already implied by the uncertainty at the core of the clinical reasoning process, but in clinical text the uncertainty is often made explicit by means of hedging. It’s therefore important to note that when we are extracting concepts from medical text, it’s very hard to make direct assertions about the real world, but we are limited to recognizing probability statements made by clinicians.

The presence class therefore captures whether a concept is present in three qualifiers. The *present* and *absent* qualifiers are used when the clinician assesses a concept as being present (or absent) with very high probability, extending beyond reasonable doubt. When neither presence nor absence is definitively asserted, the *uncertain* qualifier applies. This qualifier therefore ranges from very unlikely to very likely.

The default qualifier for presence is *present*. When the text does not indicate absence or uncertainty of a concept, we assume the writer intended to convey its presence.

## To resolve

* In future work, the *uncertain* qualifier may be further split up, for instance, into a *negative uncertain* (i.e., unlikely) qualifier and a *positive uncertain* (i.e., likely) qualifier. Or perhaps an ‘*uncertain uncertain’* qualifier in addition to those, for 50/50 cases.
* The exact threshold for absent and present should be further defined. What probability cutoff should be regarded as ‘beyond a reasonable doubt’? We can set this threshold at two standard deviations (<0.025, >0.975), but it would be even better to do some small empiric study with clinicians, to find out where each trigger term should go. Consider for example edge cases such as: very insignificant, very likely, subclinical, etc.

## Examples

trigger concept

|  |  |
| --- | --- |
| **Example** | **Qualifier** |
| Rechtszijdig fraai de middelste neusgang te visualiseren, vrij van poliepen. | *absent* |
| Tractus circulatorius: geen pijn op de borst. | *absent* |
| Een acuut coronair syndroom werd uitgesloten. | *absent* |
| Werkdiagnose maagklachten bij diclofenac gebruik en weinig intake. | *uncertain* |
| Waarschijnlijk hematurie bij reeds gepasseerde niersteen. | *uncertain* |
| Dat er toen bradypacing is geweest valt niet uit te sluiten. | *uncertain* |
| In juni 2023 longembolie waarvoor rivaroxaban met nu asymptomatische progressie. | *present* |
| PTSS en recidiverende depressie in VG. | *present* |
| Status na mild delier, heden wel slaperig. | *present* |

# Temporality

|  |  |  |
| --- | --- | --- |
| ***historical*** | ***current*** (default) | ***future*** |
| Concepts that were applicable at some point in history, but not in the last two weeks. | Concepts that were applicable in the last two weeks (potentially starting before that) up to and including the present moment. | Concepts that are potentially applicable in a future scenario. |

The temporality class places concepts in a temporal framework, ranging from past to future, relative to the document date. The *historical* and *current* qualifiers distinguish between concepts that were applicable in the past, versus concepts that are applicable in the present. The exact cutoff between *historical* and *current* is problem-specific and therefore hard to definitively establish in a general sense. In a discharge summary, everything that happened before the admission period could be considered *historical*, which can easily range up to months, while during a GP (General Practitioner) visit, events from a few days prior might be considered *historical*. For the general case, we see no reason to divert from the threshold of two weeks in the original Context paper (Harkema et al., 2009). Note that the *current* qualifier also applies when the concept is applicable in the last two weeks, but already started before that.

The *future* qualifier is applicable when a concept is described in a future scenario, for instance when describing the risk of developing a condition at a later stage, or when describing a procedure that will take place later.

## To resolve

* A way to dynamically define the threshold for *historical* and *current*, so that a cutoff can be established for each problem. In future work, we might map each concept to a timedelta (e.g., -1 year, -14 days, +5 days), but that does not fit the current qualifier framework very well. Also, it seems quite a hard problem.

## Examples

trigger concept

|  |  |
| --- | --- |
| **Example** | **Qualifier** |
| Zwanger, meerdere miskramen in de voorgeschiedenis. | *historical* |
| Progressieve autonome functiestoornissen bij eerdere dermoidcyste. | *historical* |
| Als tiener een osteotomiecorrectie beiderzijds gehad. | *historical* |
| Echocardiografisch zagen wij geen aanwijzingen voor een hypertrofe cardiomyopathie. | *current* |
| Al langer bestaande bloeddrukproblematiek. | *current* |
| CT thorax: laesie rechter onderkwab bevestigd. | *current* |
| Conservatieve maatregelen ter preventie van pulmonale infectie zijn herbesproken. | *future* |
| Mocht hij koorts en/of tachycardie ontwikkelen, dan contact opnemen met dienstdoende arts. | *future* |
| Wordt nu opgenomen middels IBS ter afwending van suïcidaliteit. | *future* |

# Experiencer

|  |  |  |
| --- | --- | --- |
| ***patient*** (default) | ***family*** | ***other*** |
| Concepts applicable to the patient related to the current document. | Concepts not applicable to the patient, but to someone with a genetic relationship to the patient. | Concepts not applicable to the patient, but to someone without a genetic relationship to the patient. |

The experiencer qualifier distinguishes between concepts that apply to the *patient*, to those that apply to *family* members with a genetic relationship to the patient, and *other* individuals with no genetic relationship to the patient (e.g. acquaintances). Clinical documents are typically obtained from electronic health records, where the relation between a document and a patient is explicit. Since a patient is a well separated entity, there is usually little ambiguity which class applies. If a concept applies to both the patient and another person, the patient label should be selected.

## Examples

trigger concept

|  |  |
| --- | --- |
| **Example** | **Qualifier** |
| Behandeling in WKZ ivm diabetes beeindigd. | *patient* |
| Pte wil geen medicatie tegen parkinson ivm slechte ervaringen broer | *patient* |
| X-enkel rechts: schuine fractuur laterale malleolus | *patient* |
| Familieanamnese omvat: ADD/ADHD: broer | *family* |
| Moederszijde: voor zover bekend geen kanker | *family* |
| 2. Covid positieve huisgenoot | *other* |

# References

Afzal, Z., Pons, E., Kang, N. *et al.* ContextD: an algorithm to identify contextual properties of medical terms in a Dutch clinical corpus. *BMC Bioinformatics* **15**, 373 (2014). <https://doi.org/10.1186/s12859-014-0373-3>

Harkema H, Dowling JN, Thornblade T, Chapman WW. ConText: an algorithm for determining negation, experiencer, and temporal status from clinical reports. J Biomed Inform. 2009 Oct;42(5):839-51. doi: 10.1016/j.jbi.2009.05.002. Epub 2009 May 10. PMID: 19435614; PMCID: PMC2757457.