Building a Norwegian Lexical Resource for Medical Entity Recognition

Ildikó Pilán*, Pål H. Brekke†, Lilja Øvrelid*
*Department of Informatics, University of Oslo, †Department of Cardiology, Oslo University Hospital Rikshospitalet
Oslo, Norway
ildikop@ifi.uio.no, pabrek@ous-hf.no, liljao@ifi.uio.no

Abstract

We present a large Norwegian lexical resource of categorized medical terms. The resource merges information from large medical databases, and contains over 77,000 unique entries, including automatically mapped terms from a Norwegian medical dictionary. We describe the methodology behind this automatic dictionary entry mapping based on keywords and suffixes and further present the results of a manual evaluation performed on a subset by a domain expert. The evaluation indicated that ca. 80% of the mappings were correct.

Keywords: lexical resource, medical terminology, Named Entity Recognition, clinical text, Norwegian

1. Introduction

Named Entity Recognition (NER) is a common task within the area of clinical Natural Language Processing (NLP) with the aim of extracting critical information such as diseases and treatments from unstructured texts (Friedman et al., 1994; Xu et al., 2010; Jagannatha and Yu, 2016).

Current neural approaches to NER typically require a large amount of annotated data for a reliable performance (Ma and Hovy, 2016; Lample et al., 2016). Distant supervision (Mintz et al., 2009), however, relaxes this constraint on the training data size thanks to the combined use of information from lexical resources, a small amount of training data and large amounts of raw data. This technique has been successfully applied also in the biomedical and clinical domain (Fries et al., 2017; Shang et al., 2018). In absence of even a small amount of annotated data, categorized lexical resources can also be used as gazetteers in rule-based approaches.

There is currently no large and freely available lexical resource with categorized entity types for Norwegian medical terms to be used for clinical NER with distant supervision. This paper presents an effort to create such a resource by collecting and merging lists of terms available from a number of other smaller and more specialized resources. We implement and describe an automatic mapping method which is applied to a dictionary containing a variety of definitions for relevant terms and present an evaluation of this mapping using both inter-resource overlap and manual evaluation performed by a domain expert. The resulting lexical resource will be made freely available.

2. Background

Medical Entity Recognition often makes use of lexical resources such as lists of disease names derived from the International Statistical Classification of Diseases and Related Health Problems (ICD) resource (World Health Organization and others, 2004) or from disease information from general resources, such as the Medical Subject Headings (Lipscomb, 2000, MeSH). There has been quite a bit of work aimed at creating semantic lexicons for use in NLP from such domain-specific resources (Johnson, 1999; Liu et al., 2012).

Automated extraction of medical entities from clinical text has been the topic of several research efforts more recently, a majority aimed at English (Xu et al., 2010; Jagannatha and Yu, 2016) and Chinese clinical text (Wu et al., 2018). For a language that is very closely related to Norwegian, Skeppstedt et al. (2014) developed and evaluated an entity detection system for Findings, Disorders and Body Parts in Swedish. In order to alleviate the need for manual annotation, distant supervision has recently been applied also to entity recognition in the medical domain for English and Chinese (Shang et al., 2018; Nooralahzadeh et al., 2019).

3. Norwegian Medical Terminology Resources

There are a number of resources which contain Norwegian medical terms that could in principle be relevant for NER.

The *Medisinsk ordbok* (MO) 'Medical Dictionary' (Nylenna, 1990) contains 23,863 Norwegian medical terms of various kinds including, among others, names of diseases and treatments, anatomical terminology as well as types of medical specialists and specialization areas. The dictionary contains synonyms and one or more definitions of these terms depending on the number of senses per entry.

Other rich sources of Norwegian medical terms and their corresponding standardized codes are available from the website of *Directoratet for e-helse* 'Norwegian Directorate for e-health'. One is the Norwegian equivalent of the 10th Revision of ICD (ICD-10). The widely-used resource lists both coarse and fine-grained codes and corresponding terms relative to diseases, symptoms and findings. Another source is the Procedure Coding Schemes list (referred to as PROC here), which includes diagnostic, medical and surgical intervention names and codes (Direktoratet for e-helse, 2020). Moreover, Laboratoriekodeverket[1] 'List of laboratory codes' (LABV) contains various substance names relevant in laboratory analyses. The web page of this list also includes a shorter list of anatomical locations, which we refer to as ALOC here. Yet another resource available from the Directorate's web site is the Norwegian equivalent of

---

[1] https://ehelse.no/kodeverk/laboratoriekodeverket

the International Classification of Primary Care (ICPC-2), which includes diagnosis terms as well as health problem and medical procedure names.

The FEST (*Forskrivnings- og ekspedisjonsstøtte*, 'Prescribing and dispensing support') database[2] contains information about all medicines and other goods that can be prescribed in Norway. FEST is a publicly available resource published by *Statens legemiddelverk* 'The Norwegian Medicines Agency'.

Rama et al. (2018) present a corpus of synthetically produced clinical statements about family history in Norwegian (here dubbed FAM-HIST). The corpus is annotated with clinical entities relating to family history, such as Family Member, Condition and Event, as well as relations between these.

4. Automatic Dictionary Entry Mapping Method

The use of dictionary definitions as a source of semantic information has been the topic of quite a bit of research in lexical semantics, from the early work of Markowitz et al. (1986) where patterns in the dictionary definitions along with suffix information gave rise to a semantic lexicon to more recent efforts to embed dictionary definitions in order to derive semantic categories for phrasal units (Hill et al., 2016).

In this work, we map entries from the MO dictionary to categories, i.e. to medical entity types. We identify 12 different types of entity categories based on previous work (Zhang and Elhadad, 2013) and the inspection of MO entries. We then implement a rule-based mapping method relying on suffixes and keywords.

4.1. Mapping Strategies

The mapping method consists of four different mapping strategies: two relying on the entries themselves and two deriving the mapped category from the definitions. One of these is suffix based, the others operate based on keywords. In what follows, we describe each of these strategies in detail.

Suffix-based mapping (*strategy SUFF*) This strategy consists of mapping an entry to a category whenever its last characters match a specific suffix. Many medical terms have Greek or Latin origin resulting in suffixes that give rather clear indications of the category of an entry. We compile a list of suffixes based on both frequently occurring suffixes in the data and an online resource[3]. We only include suffixes and endings which can be mapped to an unambiguous category in the majority of cases. The complete list used for the mapping is presented in Table 1.

Keyword-based mapping Mapping entries to keywords is primarily used to map an entry to a category based on the first noun occurring in their definition (*strategy KW-1N*). To be able to detect first nouns, definitions are tokenized and part-of-speech tagged with UDPipe (Straka et al., 2016).

Category	Suffixes
CONDITION	-agi, -algi, -algia, -blastom, -cele, -cytose, -donti, -dynia, -emi, -emia, -epsi, -ism, -isme, -ismus, -itis, -oma, -pati, -plasi, -plegi, -ruptur, -sarkom, -sis, -trofi, -temi, -toni, -tropi
DISCIPLINE	-iatri, -logi
MICROORG	-coccus, -bacillus, -bacter
PERSON	-iater, -olog
PROCEDURE	-biopsi, -grafi, -metri, -skopi, -tomi
SUBSTANCE	-cillin
TOOL	-graf, -meter, -skop

Table 1: Suffix mapping.

To create a list of keywords for the mapping, we inspect the 200 most frequent nouns in the definitions and manually map the ones with a strong indication of a single category. We complement this with other frequent nouns which can be good indicators of a category. This results in a list of 168 mapped keywords, see Table 2 for some examples.

When mapping, we require the first noun of a definition to either (i) exactly match a keyword or (ii) to contain it. The latter is only applied for keywords longer than 4 characters to avoid short sequences which might over-generate false positives (e.g. *tap* 'loss' for *katapleksi* 'cataplexy'). When checking for contained keyword, we limit the position of the keyword match to the second character onward in the first noun to approximate the occurrence of a keyword as the second part of a compound as this is more indicative of categories. Given that many dictionary entries are also compounds, we apply the mapping based on contained keyword also to the entries themselves (*strategy KW-E*).

When applying keyword-based mapping to definitions, before detecting the first noun, we remove those nouns and phrases which have little added semantic value relevant for the category. These include prepositional phrases forming a complex noun phrase typical of definitions (e.g. *form av* 'form of'), nouns not indicative of a category (e.g. *uttrykk* 'expression') and abbreviations (*plur.* 'plural', *lat.* 'Latin'). During the mapping procedure, first each strategy casts a vote on the category. In case of multiple votes with a disagreement, the category is based on a single mapping strategy chosen following a specific order, starting from the strategy with the highest expected precision and continuing with the ones with increasingly high recall as follows: SUFF → KW-E → KW-1N. After a first iteration of mapping, we perform a second iteration and map uncategorized entries if there is an entry already mapped available for the first noun in their definition (*strategy ITER*).

The MO resource contains altogether 2,387 synonyms, which were treated as separate entries with the same definition. The number of entries with multiple meanings (and definitions) were merely 360 in total, amounting to 1.5%. Since such polysemous entries were so rare, we consider only the first sense of each entry.

The methodology outlined above could be applied also for categorizing medical terminology in other languages via, for example, machine translating the list of keywords

---

Category	Description	Example keywords	Mapped entry examples
ABBREV	abbreviations, acronyms	*forkortelse* 'abbreviation'	*Ahus, ADH*
ANAT-LOC	anatomical locations	*celler* 'cells', *muskel* 'muscle', *kroppsdel* 'bodypart'	*fødselskanalen* 'birth-channel', *halsmusklene* 'throat-muscles'
CONDITION	diseases, findings	*sykdom* 'disease', *tilstand* 'condition', *mangel* 'deficiency'	*leukemi* 'leukemia', *leverkoma* 'hepatic coma'
DISCIPLINE	medical disciplines	*studium* 'study', *forskning* 'research', *teori* 'theory'	*dietetikk* 'diethetics', *biomekanikk* 'biomechanics'
MICROORG	microorganisms of different kind	*bakterie* 'bacteoria', *organisme* 'organism', *virus* 'virus'	*kolibakterie* 'colibacteria', *blodparasitter* 'blood parasites'
ORGANIZATION	institutions and organizations	*foretak* 'company', *institutt* 'institute'	*Røde Kors* 'Red Cross', *sanatorium* 'sanatorium'
PERSON	types of practitioner or patient	*lege* 'doctor', *pasient* 'patient', *individ* 'individual'	*myop* 'myope', *nevrolog* 'neurologist'
PHYSIOLOGY	physiological functions	*refleks* 'reflex', *sammentrekning* 'contraction'	*adsorpsjon* 'absorption', *forbrenning* 'burning'
PROCEDURE	procedure and treatment types	*behandling* 'treatment', *fjerning* 'removal'	*nyrebiopsi* 'kidney biopsy', *detoksifisering* 'detoxification'
SERVICE	types of services	*tjeneste* 'service', *omsorg* 'care'	*tannhelsetjeneste*, ' dental service', *sjelesorg* 'counseling'
SUBSTANCE	medicines and other substances	*stoff* 'substance', *løsning* 'solution' *medikament* 'drug'	*aspartam* 'aspartam', *paracetamol* 'paracetamol'
TOOL	instruments and tools	*instrument* 'instrument', *verktøy* 'tool'	*diatermikniv* 'diathermy blade', *defibrillator* 'defibrillator'

Table 2: List of entity type categories, keywords and mapped entries.

(or terms) used and making small language-specific orthographic adjustments to the suffix mappings from Table 1. Such suffixes are often adapted to the orthographic conventions of a certain language, as also Grigonytė et al. (2016) found in the case of Swedish.

5. Mapping Results

The results of the category mapping for MO based on the methodology outlined in Section 4. is presented in Table 3.

Category	# entries
CONDITION	5,522
SUBSTANCE	2,216
PROCEDURE	1,467
DISCIPLINE	418
ANAT-LOC	408
PERSON	282
MICROORG	227
ABBREV	216
TOOL	210
PHYSIOLOGY	132
ORGANIZATION	81
SERVICE	48
Total mapped	**11,227**
Not mapped	12,636
Total	**23,863**

Table 3: Mapping results for MO.

The percentage of mapped entries was 47%, almost half of all available entries in MO. The other terms, which were not

mapped, did not match either any of the suffixes or the keywords used. The latter includes, among others, cases where the first noun in the definition was a synonym of the term and hence too specific to be included in the list of keywords used (e.g. the term *klorose* 'chlorosis', a type of anemia occurring mosty in adolescent girls is defined using *jomfrusyk* 'virgin sick').

Based on some manual inspection, most non-mapped terms would fit one of the categories proposed, with few exceptions that might lead to rather small categories, such as regulations (e.g. *internasjonalt helsereglement* 'international health regulations'). Several non-mapped terms should belong to the ANAT-LOC category. The proposed keyword-based methods would often be ambiguous for these terms and could indicate either an anatomical location or a medical condition related to it. For example, both the ANAT-LOC *hjertekammer* 'ventricle' and the CONDITION *panserhjerte* 'armoured heart' contain the keyword *hjerte* 'heart' and have this word also as the first noun in their definition, their category could thus not be determined by our method. Additional databases containing a detailed list of anatomical location terms are therefore particularly useful for expanding our resource.

We also inspected the distribution of the mapping strategies used (see Table 4), where MULTI stands for a category selected based on the unanimous vote of multiple voting strategies. We can observe that the most frequently used strategy was KW-IN. Mappings based on multiple voting strategies selecting the same category were also rather common, occurring in 21% of all mapped entries.

Strategy	# entries
KW-1N	5,489
MULTI	2,397
ITER	1,157
SUFF	1,096
KW-E	1,088
Total	**11,227**

Table 4: Distribution of mapping strategy use.

Resource	Category	# entries
MO	Multiple	11,227
ALOC	Multiple	287
FAM-HIST	COND.	283
FEST	SUBST.	26,234
ICD-10	COND.	10,765
ICPC-2	Multiple	9,420
LABV	SUBST.	14,193
PROC	PROC.	8,883
Total	**N/A**	**81,292**

Table 5: Number and type of entries in different resources.

6. Resource Merging

The mapped MO entries were complemented with data from the other resources described in Section 3. The mapping for these resources was straightforward since each resource contained either one specific type of entity or manual annotation was available.

At a closer inspection, we found that the ALOC list contains, besides anatomic locations, several terms which could belong to more than one category depending on the context of their use, e.g. *tracheostomi* 'tracheostomy' could either be ANAT-LOC referring to the hole created during a tracheostomy or it could refer to the procedure itself. These cases were mapped to PROC for reasons of consistency with the suffix-based mapping applied, but it might be worth to accommodate multiple categories in future versions. This list has been manually revised by a medical expert who disambiguated the category consistently with the mapping methodology used.

From FAM-HIST, we collected all occurrences of condition and event entities and mapped them to our CONDITION category. The SUBSTANCE category was augmented, in part, based on the FEST resource. The terms collected from FEST included substance names (also in English, when available) as well as medical product names with and without strength information. From ICD-10, both the disease names corresponding to the 3 and the 4 digit codes were preserved. Only 16% of the ICD codes were 3 digit codes.

From ICPC-2, we included all terms, sub-terms and short forms under the CONDITION category except for the terms appearing in the *Procedure codes* chapter, which were mapped to the PROCEDURE category. Terms from the *Social problems* chapter were excluded as most of these were not strictly speaking medical conditions (e.g. *lav inntekt* 'low income'). We observed a minor difference compared to ICD between some terms associated to the same code (e.g. *Blindtarmsbetennelse* vs. *Uspesifisert appendisitt* for code K37, appendicitis).

In the case of LABV, we included under the SUBSTANCE category all substance names, medicine and other medical product names and brands together with type and strength information when available (e.g. *Kortison Tab 25 mg* 'Cortisone Tablet 25 mg'). Lastly, all codes from PROC were included without any filtering.

Table 5 presents the amount of total entries available from various resources compared to MO. The total number of categorized entries created after merging and excluding all inter-resource overlaps was 78,105 with the original casing and 77,320 when normalizing all entries to lowercase.

7. Resource-based Automatic Evaluation

Thanks to a certain amount of overlap between the mapped MO entries and the other resources, we can use information from the latter to automatically evaluate the former. Table 6 shows the overlap and the percentage of correct mappings.

Resource	# overlap	Correct (%)	Category
ALOC	33	57.6	Multiple
FAM-HIST	22	63.6	COND.
FEST	744	97.3	SUBST.
ICD-10	307	97.7	COND.
ICPC-2	886	94.0	Multiple
LABV	297	85.5	SUBST.
PROC	89	97.8	PROC.

Table 6: Evaluation results of the mapped MO entries.

On average, 85% mappings were correct out of the total of 2,378 overlapping terms from the resources listed in Table 6. Approximately 21% of all mapped terms from MO were thus evaluated (and corrected) automatically with the help of the other resources. Most misclassifications occurred with the ALOC and FAM-HIST resources and concerned the ANAT-LOC and CONDITION categories.

8. Manual Evaluation

Given that the overlap between MO and the other resources was limited to certain categories, we further performed a manual evaluation of the automatically mapped MO entries in order to assess their quality.

We randomly selected 1,128 terms to evaluate manually, aiming at a balanced amount per category (100 each) and mapping method. We included all available terms for categories where the total amount of terms remained below 100. The terms were categorized by a medical expert without access to the automatically mapped categories and the mapping method used. We present the per-category precision and recall in Table 7, where the number of terms in the last column refers to manually assigned labels.

112 terms were labeled as 'OTHER' in cases where a term did not belong to any of the 12 categories indicated or when terms were outside of the area of expertise of the evaluator. Table 7 excludes OTHER, as this was not part of the automatically mapped categories. The percentage of correctly categorized entries including and excluding terms la-

Category	Prec	Recall	#
ABBREV	0.969	0.750	124
ANAT-LOC	0.928	0.796	113
CONDITION	0.915	0.623	138
DISCIPLINE	0.702	0.855	69
MICROORG	0.871	0.976	83
ORGANIZATION	0.548	0.714	56
PERSON	0.593	0.923	52
PHYSIOLOGY	0.710	0.815	81
PROCEDURE	0.793	0.821	84
SERVICE	0.667	0.468	47
SUBSTANCE	0.809	0.905	84
TOOL	0.846	0.906	85
Total	**0.779**	**0.796**	**1,016**
ORG+SER	0.830	0.854	103
Total ORG+SER	**0.815**	**0.839**	**1,016**

Table 7: Manual evaluation results.

beled as OTHER, was 71.5% and 79.4% respectively. In 20 cases, SERVICE and ORGANIZATION were indicated as alternative labels to each other. We therefore compute evaluation measures also with these two categories merged (ORG+SER). This yields in total 82% correct labels when excluding OTHER.

According to the confusion matrix in Figure 1, most automatic categorization errors occurred between CONDITION and PHYSIOLOGY. (SERVICE was mapped to ORGANIZATION here.)
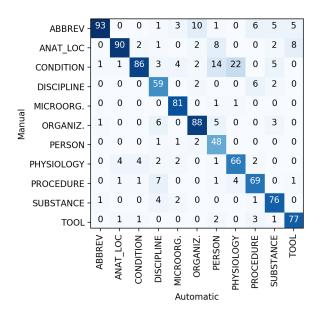

Figure 1: Confusion matrix over categories.

Errors related to the PERSON category were mostly connected to the use of *person* as keyword with the KW-E strategy, which generated false positives such as *schizoid personlighetstype* 'schizoid personality type'. Some categorization errors occurred because of the lack of prefix information, e.g. in the case of the keyword *refleks* 'reflex' in *arefleksi* 'areflexia' and *hyperrefleksi* 'hyper-

reflexia', which were both mapped to PHYSIOLOGY instead of CONDITION. This indicates that taking into consideration prefixes would contribute to improving the automatic categorization, especially for the KW-E strategy. The category label confusions between TOOL and ANAT-LOC originated from the keyword *apparat*, which proved to be ambiguous for the proposed categories, not only meaning 'device' and thus mappable to TOOL, but also meaning 'apparatus, system' as in *immunapparatet* 'immune system' and thus belonging to ANAT-LOC.

Most correct mappings (88.3%) with a single strategy were obtained using suffixed (SUFF), followed by the keyword mapping from first nouns (KW-1N, 79.9%) and entries (KW-E, 76.6%). The iterative mapping (ITER) yielded considerably fewer correct mappings, only 64.7%. When multiple strategies opted for the same category label, 98.2% of terms were correctly categorized.

As a final step during the resource creation, we revised the automatic categories based on the manually assigned ones. The updated count of terms per category in the resource after merging with other databases (eliminating overlap) and incorporating the evaluation results is reported in Table 8.

Category	# entries
SUBSTANCE	41,365
CONDITION	24,071
PROCEDURE	10,420
ANAT-LOC	658
DISCIPLINE	387
ABBREV	236
PERSON	232
TOOL	216
MICROORGANISM	193
OTHER	112
PHYSIOLOGY	112
ORGANIZATION	103
Total (original casing)	**78,105**

Table 8: Final term counts per category in the resource.

9. Conclusion

We introduced the first Norwegian lexical resource of categorized medical entities and provided an overview of the process of its creation. The resource unites information from medical databases as well as entries automatically mapped from a medical lexicon. A manual evaluation of a subset of the mapped terms confirmed that the automatic mappings were of a suitable quality to be used as additional supervision signal with machine learning based NER approaches. In future work we plan to apply the resource in medical entity recognition for Norwegian, using it to provide initial categories for distant supervision. We also plan to perform annotations with multiple raters and measure inter-annotator agreement for the proposed categories.

10. Acknowledgments

11. Bibliographical References

Direktoratet for e-helse. (2020). Prosedyrekodeverkene 'Procedure Coding Schemes'. `https://ehelse.no/kodeverk/prosedyrekodeverkene-kodeverk-for-medisinske-kirurgiske-og-radiologiske-prosedyrer-ncmp-ncsp-og-ncrp`. Accessed: 2020-02-10.

Friedman, C., Alderson, P. O., Austin, J. H., Cimino, J. J., and Johnson, S. B. (1994). A general natural-language text processor for clinical radiology. *Journal of the American Medical Informatics Association*, 1:161 – 174.

Fries, J. A., Wu, S., Ratner, A., and Ré, C. (2017). Swellshark: A generative model for biomedical named entity recognition without labeled data. *CoRR*, abs/1704.06360.

Grigonytė, G., Kvist, M., Wirén, M., Velupillai, S., and Henriksson, A. (2016). Swedification patterns of Latin and Greek affixes in clinical text. *Nordic Journal of Linguistics*, 39(1):5–37.

Hill, F., Cho, K., Korhonen, A., and Bengio, Y. (2016). Learning to understand phrases by embedding the dictionary. *Transactions of the Association for Computational Linguistics*, 4:17–30.

Jagannatha, A. N. and Yu, H. (2016). Bidirectional RNN for medical event detection in electronic health records. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 473–482, San Diego, California, June. Association for Computational Linguistics.

Johnson, S. B. (1999). A semantic lexicon for medical language processing. *Journal of the American Medical Informatics Association*, 6(3):205–218.

Lample, G., Ballesteros, M., Subramanian, S., Kawakami, K., and Dyer, C. (2016). Neural architectures for named entity recognition. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 260–270, San Diego, California, June. Association for Computational Linguistics.

Lipscomb, C. E. (2000). Medical subject headings (MeSH). *Bulletin of the Medical Library Association*, 88(3):265–266.

Liu, H., Wu, S. T., Li, D., Jonnalagadda, S., Sohn, S., Wagholikar, K., Haug, P. J., Huff, S. M., and Chute, C. G. (2012). Towards a semantic lexicon for clinical natural language processing. In *AMIA Annual Symposium Proceedings*, volume 2012, page 568. American Medical Informatics Association.

Ma, X. and Hovy, E. (2016). End-to-end sequence labeling via bi-directional LSTM-CNNs-CRF. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1064–1074, Berlin, Germany, August. Association for Computational Linguistics.

Markowitz, J., Ahlswede, T., and Evens, M. (1986). Semantically significant patterns in dictionary definitions. In *Proceedings of the 24th annual meeting on Association for Computational Linguistics*, pages 112–119. Association for Computational Linguistics.

Mintz, M., Bills, S., Snow, R., and Jurafsky, D. (2009). Distant supervision for relation extraction without labeled data. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 2 - Volume 2*, ACL '09, pages 1003–1011, Stroudsburg, PA, USA. Association for Computational Linguistics.

Nooralahzadeh, F., Lønning, J. T., and Øvrelid, L. (2019). Reinforcement-based denoising of distantly supervised NER with partial annotation. In *Proceedings of the 2nd Workshop on Deep Learning Approaches for Low-Resource NLP (DeepLo 2019)*, pages 225–233, Hong Kong, China, November. Association for Computational Linguistics.

Nylenna, M. (1990). *Medisinsk Ordbok*. Kunnskapsforlaget.

Rama, T., Brekke, P., Nytrø, Ø., and Øvrelid, L. (2018). Iterative development of family history annotation guidelines using a synthetic corpus of clinical text. In *Proceedings of the 9th International Workshop on Health Text Mining and Information Analysis (LOUHI 2018)*.

Shang, J., Liu, L., Ren, X., Gu, X., Ren, T., and Han, J. (2018). Learning named entity tagger using domain-specific dictionary. In *Proceedings of EMNLP*.

Skeppstedt, M., Kvist, M., Nilsson, G. H., and Dalianis, H. (2014). Automatic recognition of disorders, findings, pharmaceuticals and body structures from clinical text: An annotation and machine learning study. *Journal of Biomedical Informatics*, 49:148 – 158.

Straka, M., Hajic, J., and Straková, J. (2016). UDPipe: trainable pipeline for processing CoNLL-U files performing tokenization, morphological analysis, POS tagging and parsing. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pages 4290–4297.

World Health Organization et al. (2004). ICD-10: International statistical classification of diseases and related health problems: Tenth revision.

Wu, Y., Jiang, M., Xu, J., Zhi, D., and Xu, H. (2018). Clinical named entity recognition using deep learning models. *AMIA ... Annual Symposium proceedings. AMIA Symposium*, 2017:1812–1819, 04.

Xu, H., Stenner, S. P., Doan, S., Johnson, K. B., Waitman, L. R., and Denny, J. C. (2010). MedEx: a medication information extraction system for clinical narratives. *Journal of the American Medical Informatics Association*, 17:19 – 24.

Zhang, S. and Elhadad, N. (2013). Unsupervised biomedical named entity recognition. *J. of Biomedical Informatics*, 46(6):1088–1098.