Initial Experiments of Medication Event Extraction Using Frame Semantics

Dimitrios Kokkinakis

Centre for Language Technology; the Swedish Language Bank
Department of Swedish Language, University of Gothenburg, Gothenburg, Sweden

Abstract

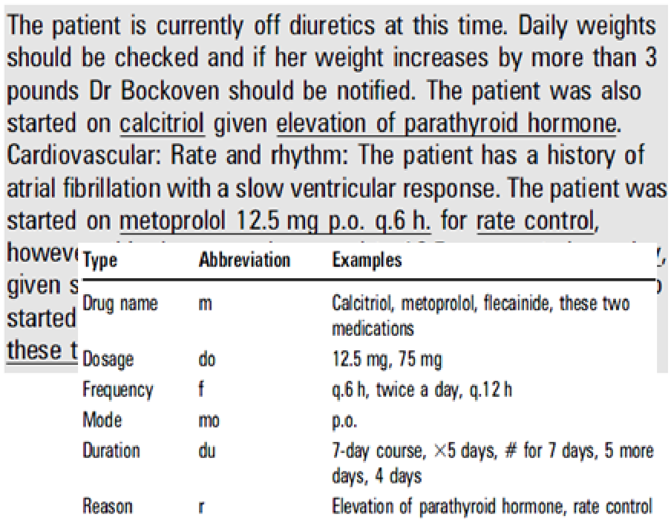
Semantic annotation of text corpora for mining complex relations and events has gained a considerable growing attention in the medical domain. The goal of this paper is to present a snapshot of ongoing work that aims to develop and apply an appropriate infrastructure for automatic event labelling and extraction in the Swedish medical domain. Annotated text samples, appropriate lexical resources (e.g. term lists and the Swedish Frame-Net++) and hybrid techniques are currently developed in order to alleviate some of the difficulties of the task. As a case study this paper presents a pilot approach based on the application of the theory of frame semantics to automatically identify and extract detailed medication information from medical texts. Medication information is often written in narrative form (e.g. in clinical records) and is therefore difficult to be acquired and used in computerized systems (e.g. decision support). Currently our approach uses a combination of generic entity and terminology taggers, specifically designed medical frames and various frame-related patterns. Future work intends to improve and enhance current results by using more annotated samples, more medically-relevant frames and combination of supervised learning techniques with the regular expression patterns.

Keywords:

Natural Language Processing; Automatic Data Processing; Semantics; Medication Extraction; Chemicals and Drugs; Dosage.

Introduction

Semantic annotation of text corpora for mining complex relations and events is a challenging research topic that has gained a considerable growing attention in the medical domain [1,2]. The goal of this paper is to present a snapshot of ongoing work that aims to develop and apply the appropriate infrastructure for automatic event labelling and extraction in the Swedish medical domain. Our approach is closely related to information extraction (IE), a technology that has a direct correlation with frame-like structures as described in the FrameNet (see below and also Appendix A). Templates in the context of IE are frame-like structures with slots representing event information. Most event-based IE approaches are designed to identify role fillers that appear as arguments to event verbs or nouns, either explicitly via syntactic relations or implicitly via proximity. As a case study, the paper presents an approach to automatically extract detailed medication information/events from medical texts (scientific papers, clinical records etc.) based on the theory of frame semantics. Medication information is often written in a narrative form and is therefore difficult to use in computerized systems or acquired by mining technologies.



The patient is currently off diuretics at this time. Daily weights should be checked and if her weight increases by more than 3 pounds Dr Bockoven should be notified. The patient was also started on calcitriol given elevation of parathyroid hormone. Cardiovascular: Rate and rhythm: The patient has a history of atrial fibrillation with a slow ventricular response. The patient was started on metoprolol 12.5 mg p.o. q.6 h. for rate control,
howev
given s
started
these t

Type	Abbreviation	Examples
Drug name	m	Calcitriol, metoprolol, flecainide, these two medications
Dosage	do	12.5 mg, 75 mg
Frequency	f	q.6 h, twice a day, q.12 h
Mode	mo	p.o.
Duration	du	7-day course, ×5 days, # for 7 days, 5 more days, 4 days
Reason	r	Elevation of parathyroid hormone, rate control

Figure 1. Example of discharge summary (left) and extracted information on drugs (bottom-right); figures taken from [7].

Our approach uses a combination of adapted entity and terminology taggers, specifically designed medical frames and various frame related regular expression patterns.

Several systems for medication event extraction have been reported in the last couple of years [9]. However, the most relevant research initiatives with respect to the presented work is the third i2b2 Workshop on NLP Challenges for Clinical Records (designed as an information extraction task; see <https://www.i2b2.org/NLP/Medication/>) which focused on the extraction of medications and medication-related information from discharge summaries [4-8]. Medication extraction has numerous applications in e.g. pharmacovigilance, pharmacogenetics research, medication surveillance, clinical decision support, biomedical research and as an input or pre-processing step for medical language processing tools intended for data mining and knowledge discovery. Figure 1, taken from [7] shows an example of underlined medication-related information from narrative discharge summaries (left) and how structured extraction of the targeted categories of information on drugs might look like (right). In the near future we intend to improve and enhance the extracted results using (larger) manually annotated samples and supervised learning in a much larger scale than we have been able to do so far [3] and particularly for all other types of medical-related frames.

Theoretical Background

The FrameNet approach is based on the linguistic theory of frame semantics [10] supported by corpus evidence. A semantic frame is a script-like structure of concepts, which are linked to the meanings of linguistic units and associated with a specific event or state. Each frame identifies a set of frame elements, which are frame specific semantic roles; both so called core roles, *arguments*, tightly coupled with the particular meaning of the frame and more generic non-core ones, *adjuncts* or *modifiers* which to large extent are event-independent semantic roles. Furthermore, roles may be expressed overtly, left unexpressed or not explicitly linked to the frame via linguistic conventions (null instantiations). Here, we only deal with the first type of such roles. FN documents the range of semantic and syntactic combinatory possibilities of frame evoking lexical units (LU), phrases and clauses by abstracting away from syntactic differences. A LU can evoke a frame, and its syntactic dependents can fill the frame element slots. Since a LU is the pairing of a word with a meaning, each sense of a polysemous word belongs to a different semantic frame, Moreover, since a single frame element can have different grammatical realizations it can enhance the investigation of combinatorial possibilities more precisely than other standard lexical resources such as WordNet.

Materials and Methods

The Swedish FrameNet

This paper deals with the "Administration_of_medication"-frame, part of the Swedish FrameNet (SweFN). SweFN is a lexical resource under development, based on the English version of FrameNet constructed by the Berkeley research group. The SweFN is available as a free resource and its latest version can be found here: <http://spraakbanken.gu.se/swe/forskning/swefn/>. The SweFN frames and frame names correspond to the English ones, with some exceptions, as to the selection of frame elements including definitions and internal relations. The meta-information about the frames, such as semantic relations between frames, is also transferred from the Berkeley FrameNet. Compared to the Berkeley FrameNet, SweFN is expanded with information about the domain of the frames, at present: general language, the medical and the art domain. The frames also contain notation about semantic types. FN facilitates modelling the mapping of form and meaning within these structures in the medical discourse through manual annotation of example sentences and automatic summarization of the resulting annotations.

Since frame classification is based on general-domain frame semantics, several efforts have been described to domain adaptations [11]. For instance, the Cure frame describes a situation involving a number of core roles such as: Affliction, Healer, Medication, Patient etc., and is evoked by lexical units such as *detoxify, heal, cure, surgery, treat, recover,* etc. The word in bold face below evokes the Cure frame: "[Steloperation av fotleden$_{\text{TREATMENT}}$] **lindrar** [smärta$_{\text{AFFLICTION}}$] [väl$_{\text{MANNER}}$] men medför en del komplikationer" (litt. 'Lumbar fusion operation of the ankle reduces pain well, but entails some complications'). Medical frames in SweFN include: Administration_of_medication; Cure; Addiction; Recovery; Experience_bodily_harm; Falling_ill; People_by_disease etc.

Relevant resources

The following resources (textual, terminological, etc.) have been used for both extracting relevant text samples and also aiding the recognition of relevant frame elements in the samples:

- All text samples are taken from the MEDLEX corpus [12] using key word searches and selecting random sentences.
- The FASS, which is the Swedish national formulary: FASS contains a list of medicines that are approved for prescription throughout Sweden. The FASS version of November 2011, which contains over 12.500 names of registered drugs, was used in this work. For example:

tradename/substance	form	atc
Alvedon forte	Filmdragerad tablett	N02BE01

- The Swedish SNOMED CT's hierarchies, particularly the *substance* hierarchy (including *products*) which contains "concepts that can be used for recording active chemical constituents of drug projects, food and chemical allergens, adverse reactions, toxicity or poisoning information, and physicians and nursing orders" <http://www.ihtsdo.org/snomed-ct/snomed-ct0/snomed-ct-hierarchies/substance/>. Moreover, the Administration_of_Medication frame states that, if possible, the Purpose of the medication is also required to be identified, usually a disease; in the same manner also Body_System is required to be identified if explicitly stated in a medication event context), Examples of SNOMED CT:

term	hierarchy	concept-id
kadaverin	subst.	68837002
Azitromycin	product	96034006

- The Swedish MeSH's category D, Chemicals and Drugs which contains ca 5.900 terms. MeSH is a good source of synonymous terms, since the Swedish SNOMED CT contains no synonyms.

term	MeSH hierarchy
Paracetamol	D02.065.199.092.040
Antibiotika	D27.505.954.122.085

- Semi-automatic acquired drug/substance lexicon extensions (e.g. generic expressions of drugs, misspellings) and a few relevant affixes as well as names of drug classes/types than drug names, such as antibiotica, antipsychotics, antidepressants, steroids, analgesics, anticoagulants, contraceptives, hormones etc.

  artotec vs artrotec, *b-blockerare* vs beta-blockerare, *diklofinak* vs diklofenak

- List of drug forms, such as: pill, tablet, capsule, liquid solution, suspension, injection, cream, gel, liniment, lotion, ointment, suppository etc.
- List of drug administration paths, such as: intranasal, intravenous, intravesical, subcutaneous etc.
- List of volume units such as: g, mg, milligram, mcg, mikrogram, MU, mmol, millimol, ml etc.
- List of relevant abbreviations and variants commonly used in prescription of drugs (e.g. shortened drug administration paths) and other drug-related information, such as:

  iv, i.v., im, i.m. sc, s.c., po, p.o., vb, v.b., V b, T, inj., tbl, …

Method

We manually annotated a sample of 106 sentences with all possible frame elements [14] and *cf.* Appendix B. Administration of medication events were also annotates for negation and speculation. In appendix B there is an example of such construction (the first example shown is negated, i.e. crossed annotation on the interface). Through manual analysis of the annotated examples we get an in-depth understanding of how medication events can be actually expressed in real data in order to be able to model rules (regular expressions) to be used for the task and also have available training data for future planned supervised learning extensions. Annotated example sentences are shown below, the XML-like labels should be self-explanatory, see also table 1 and Appendix (B).

We generally followed similar methodology described in [5,7,9]. We applied an existing Swedish named entity recognizer (NER) which (among other entities) identifies and annotates time expressions (marked as TIMEX) as well as various types of numerical information (marked as NUMEX) with appropriate labels in the attributes TYPE and SuBType (*cf.* [13]). These annotations are important since they are both required by the frame and are very common in the context of medication event expressions. The following example illustrates how the NER-tagger annotates occurrences of time ("TIMEX/TME"); frequency ("NUMEX/FRQ") and dosage ("NUMEX/DSG"): Åtta patienter erhöll Recormon före operationen, i dosering 2 000 IE subkutant tre gånger per vecka under tre veckor (litt. 'Eight patients received Recormon before surgery, dosage 2000 IU subcutaneously three times per week for three weeks') is annotated by the NER as "Åtta patienter erhöll Recormon före operationen , i dosering *<NUMEX TYPE="MSR" SBT="DSG">*2 000 IE*</NUMEX>* subkutant *<NUMEX TYPE="MSR" SBT="FRQ">tre gånger per vecka </NUMEX> <TIMEX TYPE="TME" SBT="DAT">under tre veckor</TIMEX>*"

In total, 136 Administration_of_Medication events could be found in the whole sample. As expected, Drug_name (the trigger element) and Dosage had most occurrences in the whole material, 100 and 66 respectively. These sentences were used for making the rules described below. All annotated samples are available from: <http://demo.spraakdata.gu.se/brat/#/sweFNAdminOfMed_dk/adminOfMedication>.

- <Drug_name>Kåvepenin</Drug_name> <Drug_strength>1 g</Drug_strength> <Frequency>x 3 </Frequency> och <Drug_name>Mollipect </Drug_name>. (litt. 'Kåvepenin 1 g X 3 and Mollipect.')
- Lugnande besked , rec <Drug_name>Tradil</Drug_name> <Drug_strength>400 mg</Drug_strength> <Frequency>1 x 1-2 </Frequency> (litt. 'Reassurance, rec Seractil 400 mg 1 x 1-2')
- Förnyat E-rec <Drug_form>T</Drug_form> <Drug_name>Stesolid</Drug_name> <Drug_ strength>5 mg </Drug_strength> <Frequency>½ x 4</Frequency>, 100 st 4 uttag <Duration>1 mån intervall</Duration> (litt. 'Renewed E-rec T Valium 5 mg x 4 ½, 100 pcs 4 during 1 month intervals')
- <Drug_name>Ventoline Diskus</Drug_name> <Drug_strength>0,2 mg</Drug_strength> <Time_of_the_day>morgon , kl 11</Time_of_ the_day> , <Time_of_the_day>kvällen</Time _of_the_day> samt ibland <Time_of_the_day>nattetid</Time_of_the_day>. (litt. 'Ventolin Diskus 0.2 mg morning, 11 o'clock, evening and sometimes at night.')

Table 1 - Frame Elements of the "Administration_of_Medication" frame

FrameElement	Definition / Example
Drug_form	pill, tablet, capsule, injection, gel, ...
Drug_name	Abboticin, Aberela, Abilify, Absenor, ...
Drug_strength	strength of the active ingredient(s)
Body_system	the part or area of the body affected by the administration, condition or disease
Dosage	the amount of a single medication used in each administration
Drug_type	a very general or generic name of a substance, such as antidepressants, steroids, analgesics, anticoagulants, hormones,...
Duration	the amount of time for which the activity / the medication is to be administered
Frequency	the number (or description of the number) of times a medication takes place over a period of time
Manner	usually various adverbials, such as slowly, fast, p.r.n. (Latin "pro re nata"), periodically, ...
Purpose	the medical reason for which the medication is stated to be given (often) a medical problem
Circumstance	the existing conditions or state of affairs surrounding and affecting an event
Route_of_drug_administration	path by which a drug, poison, or other substance is taken into the body, i.e. intranasal, intravenous, intravesical, ...
Time Time_of_the_day Time_elapsed_between_administration	various time related frame elements

The following steps have been used for the automatic annotation of the frame elements of the `Administration_of_medication` frame. These steps are applied at the sentence level only:

- start by identifying and annotating drug names (e.g. FASS) or drug name classes according to SNOMED CT or MeSH. These drug names are used as triggers for continuing the following processing steps.
- the NER-tagger is then applied and potential time, frequency or other important numerical entities are identified and annotated.
- SNOMED CT is applied once again in order to identify potential `Purpose` (e.g. Disorders) and `Body_system`.
- lexical rules based on lists of drug forms, administration path etc., implemented as regular expressions are applied for the recognition and annotation of relevant frame elements . It has been hypothesized that medication-related information is most often found in the portion of text following a drug name [5] --- in the sample we have observed some similar patterns that we also try to model in the rules, such as the most frequent: "`<Drug_name> <Drug_strength> <Frequency>`" (10 occurrences).
- in the same manner as previously, we prepared some data (40 sentences annotated and evaluated by the author) for testing the different steps. Nevertheless, it would have been advantageous if (trained) experts, e.g. physicians, could annotate the test data but that was prohibitive at the moment but will be considered in the future.
- normalizing the labels from the different processes, processing and scoring of the results, i.e. precision and recall for each frame element is calculated.

Results

Table 2 shows the evaluation results for each frame element that had more than 10 occurrences in the test data. Here *Precision* measures the amount of elements correctly labeled out of the total number of all elements labeled by the rules; while *Recall* measures the amount of elements correctly labeled given all of the elements in the sample. The evaluation results are based on 40 sentences that were annotated separately from the annotated sample used for the creation of the pattern matching rules.

Some of the frame elements could not be found in the limited test sample, while some had very few occurrences and we chose not to formally evaluate at this stage of the processing, for instance `Body_system`. In the test sample the most frequent elements were `Drug_name` with 53 occurrences and `Drug_strength` with 36. The vertical level evaluation assess the extraction of each frame element individually. So far, the extraction of `Purpose` and `Circumstance` seem the most problematic since these elements shows great common language variability, such as long phrases, vague expressions, acronyms etc. For instance: `<Circumstance>Vid klart skyldig blindtarmsinflammation av varierande grad upp till kraftigare inflammation med tecken på vävnadsdöd i blindtarmen</Circumstance> administreras antibiotika Tienam 0,5 g x 3` (litt. '*In clear-cut case appendicitis of varying degree up to stronger inflammation with signs of necrosis in the cecum* antibiotics Tienam 0.5 g x 3 is administered') and `Om misstanke att <Purpose>ITP</Purpose> föreligger, ge Prednisolon [...]` (litt: 'If suspicion of *ITP* exists give Prednisolon'). Therefore using a simple pattern matching approach is rather insufficient for most of the cases encountered so far.

Table 2 - Evaluation results (# occurrences in the test sample)

Frame Element	#	Precision	Recall
Drug_name	53	93.8%	83.6%
Drug_strength	36	94.4%	85%
Route_of_drug_ administration	27	100%	96.2%
Dosage	21	88.2%	80.9%
Frequency	19	84.2%	80%
Drug_form	19	100%	94.4%

Another problematic aspect is observed on many cases where there is an ellipsis, that is clauses where an overt trigger word (often a verbal predicate belonging to the frame) is missing. In appendix B there is an example of such construction (the second example shown in which the given event lacks an overt trigger; i.e. `samt med kinidin tabletter` (litt. '[...] and with kinidin tablets').

Conclusions

We have outlined an approach to medication event extraction using frame semantics. Extraction of event information is a hot topic in medical research [15,16]. The driving force for the experiments is the theory of frame semantics, which allows us to work with a more holistic and detailed semantic event description than it has been previously reported in similar tasks or in efforts using for instance most traditional methods based on binary relation extraction approaches. Moreover, event extraction is more complicated and challenging than relation extraction since events usually have internal structure involving several entities as participants allowing a detailed representation of more complex statements.

Preliminary results suggest that SweFN++ seems a good start for annotating corpora. The role set described is general enough to capture a wide range of phenomena that characterize the majority of semantic arguments of medical events. The motivation for the technique used so far is simplicity. There are a number of improvements and extensions envisaged for the near future. For instance, we intend to annotate more sentence samples, and in particular on different frames. Therefore, as an immediate next step we intend to apply supervised learning as a complement to the current approach which is based on pattern matching using regular expression patterns: Patterns are manually produced and their adaptation requires time and expensive human resources so the application of supervised learning methods are considered the most appropriate alternative technique that can be used, most probably in a close interplay with regular expression patterns that are suitable for matching with high accuracy short numerical and abbreviated patterns, such as dosage. Finally, a further goal we have in mind with our current work is to use such an approach for the extraction of medication events from clinical texts i.e. medical records, where a patient's medical treatment plays an important role during the patient's medical diagnosis, treatment, etc. a critical piece of information for healthcare safety and quality.

Acknowledgments

The research presented here was supported by the Swedish Research Council (the project Swedish Framenet++, VR dnr 2010-6013) and by the University of Gothenburg through its support of the Centre for Language Technology and Språkbanken (the Swedish Language Bank).

References

[1] Ananiadou S, Pyysalo S, Tsujii J, Kell DB. Event extraction for systems biology by text mining the literature. Trends Biotechnol. 2010 Jul;28(7):381-90.

[2] Miwa M, Thompson P, McNaught J, Kell DB, and Ananiadou S. Extracting semantically enriched events from biomedical literature. BMC Bioinformatics 2012, 13:108.

[3] Johansson R, Friberg Heppin K, and Dimitrios Kokkinakis. Semantic Role Labeling with the Swedish FrameNet. Proceedings of the 8th International Conf on Language Resources and Evaluation (LREC'12), 2012, pp. 3697–3700. Istanbul, Turkey.

[4] Patrick J, Li M. High accuracy information extraction of medication information from clinical notes: 2009 i2b2 medication extraction challenge. J Am Med Inf Assoc 2010; 17:524e527.

[5] Deléger L, Grouin C, and Zweigenbaum P. Extracting medical information from narrative patient records: the case of medication-related information. J Am Med Inf Assoc 2010; 17:555e558.

[6] Doan S, Bastarache L, Klimkowski S, Denny JC, and Xu H. Integrating existing natural language processing tools for medication extraction from discharge summaries. J Am Med Inf Assoc 2010; 17:528e531.

[7] Hamon T, and Grabar N. Linguistic approach for identification of medication names and related information in clinical narratives. J Am Med Inf Assoc 2010; 17:549e554.

[8] Halgrim JR, Xia F, Solti I, Cadag E, and Uzuner Ö. A cascade of classifiers for extracting medication information from discharge summaries. J of Biomed Sem 2011, 2(Suppl 3):S2

[9] Gold S, Elhadad N, Zhu X, Cimino JJ, and Hripcsak G. Extracting Structured Medication Event Information from Discharge Summaries. AMIA Annu Symp Proc. 2008; 237–241.

[10] Fillmore CJ, Johnson CR, and Petruck MRL. Background to FrameNet. Journal of Lexicography, 2003, 16(3). <http://framenet.icsi.berkeley.edu>.

[11] Dolbey A, Ellsworth M, and Scheffczyk J. BioFrameNet: A Domain-specific FrameNet Extension with Links to Biomedical Ontologies. KR-MED: "Biomedical Ontology in Action". 2006, Baltimore, Maryland, USA.

[12] Kokkinakis D. A Semantically Annotated Swedish Medical Corpus. Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC). 2008. Marrakech, Morocco.

[13] Kokkinakis D. Reducing the Effect of Name Explosion. Proceedings of the LREC Workshop: Beyond Named Entity Recognition, Semantic labelling for NLP tasks. Fourth Language Resources and Evaluation Conference (LREC). Pp. 1-6. 2004. Lissabon, Portugal.

[14] Stenetorp P., Pyysalo S., Topic G., Ohta T., Ananiadou S. and Tsujii J. BRAT: a Web-based Tool for NLP-Assisted Text Annotation. Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics. 2012. Avignon, France.

[15] Meystre S, Savova G, Kipper-Schuler Karin C, and Hurdle JE. Extracting Information from Textual Documents in the Electronic Health Record: A Review of Recent Research. IMIA Yearbook of Medical Informatics 2008. Methods Inf Med 2008; 47 Suppl 1: 128-144.

[16] Björne J, Ginter F, Pyysalo S, Tsujii J, and Salakoski T. Complex event extraction at PubMed scale. BioInformatics Vol. 26 ISMB 2010, pages i382–i390

Address for correspondence

Dimitrios Kokkinakis. Centre for Language Technology; the Swedish Language Bank. Department of Swedish Language. Box 200. University of Gothenburg. SE-405 30. Gothenburg, Sweden. Email: dimitrios.kokkinakis@svenska.gu.se.

Appendix A

The "Administration_of_medication"-frame taken from <http://spraakbanken.gu.se/swe/forskning/swefn/utvecklingsversion> (version of September 2012). In the picture below, the left column is provided in Swedish (*ram*: frame; *domän*: domain; *semantisk typ*: semantic type; *kärnelement*: core elements; *periferielement*: non-core elements; *exempel*: examples; *sms*: compound construction; *sms-exempel*: examples of compound constructions; *saldo*: relevant lexical units taken from the Swedish lexical resource 'saldo').

Administration_of_medication

ram	**Administration_of_medication**
domän	Med
semantisk typ	
kärnelement	**Drug_form**, **Drug_name**, **Drug_strength**
periferielement	**Body_system**, **Circumstances**, **Depictive**, **Dosage**, **Drug_type**, **Duration**, **Frequency**, **Manner**, **Purpose**, **Route_of_drug_administration**, **Time**, **Time_of_the_day**, **Time_elapsed_between_administration**
exempel	Behandlas med **T Vibramycin 200 mg x 1 i 14 dagar**. **T Sumatriptan 50 mg œ-1**. **T. Suscard 2,5 mg kl. 13.10** **Tramadol 50 mg 1 tabl 2-3 ggr dagl** **Tegretol 20 mg/ml**, **30 ml x 1** **Inj Diamox 250 mg x 1 iv** **inj Fluarix 0,5 ml im i höger arm** **Benzatinpenicillin 2,4 milj IE im dag 1** Åtta patienter erhöll **Recormon före operationen**, **i dosering 2 000 IE subkutant tre gånger per vecka under tre veckor** Normalt ska en **salva** eller **kräm strykas på tunt**. Modern hade under graviditeten medicinerat med **klomipramin 50 mg dagligen**. Barnets moder hade under graviditeten medicinerat med **klomipramin 60 mg x 1**. Modern hade under graviditeten medicinerat med **klomipramin 75 mg /dag**. **Läkemedlet** kan **administreras peroralt**, **rektalt** och **på huden som gel**.
sms	
sms-exempel	
saldo	**nn:** applicering..1 drog..1 glidkräm..1 glidsalva..1 herpessalva..1 hudkräm..1 injektion..1 injicerande..1 injicering..1 insprutning..1 kräm..1 läkemedel..1 medicin..2 piller..1 påläggande..2 påläggning..2 salva..1 tablett..1 ögondropp..1 ögondroppar..1 **vb:** applicera..2 gnida..1 injicera..1 **vbm:** droppa_in..2 lägga_på..1 stryka_på..1

Appendix B

Part of the annotated data with the frame elements of the `Administration_of_Medication` frame. The BRAT system [14] has been used for the manual annotation of the frame elements.

The first example shown below is actually annotated as "Negated" (crossed annotation) i.e. `bör undvikas` (litt. 'ought to be avoided'). The second example illustrates an elliptic annotation, where the second event lacks an overt trigger; i.e. [...] `samt med kinidin tabletter` (litt. '[...] and with kinidin tablets').