NTNU

Norwegian University of
Science and Technology

Department of Computer and Information Science

TDT4215 — Web-Intelligence

---

# Project Report — Group One

---

*Authors:*
Even Wiik Thomassen
Terje Snarby
Weilin Wang

Word count: 5300

March 29, 2012

**Abstract**

Classifying treatment to patients can be complicated and error-prone. It could be beneficial to provide a system to health professionals that can automatically classify patient notes. We have created such a system with the help of vector space model based on TF-IDF and probabilistic model based on BM25F. This paper describes and evaluates such a system.

# CONTENTS

# LIST OF TABLES

# CHAPTER 1
# INTRODUCTION

This paper documents the mandatory group assignment in TDT4215 Web-intelligence at Norwegian University of Science and Technology (NTNU).

The assignment asks us to implement a search system that utilizes patient record notes as a search query to therapy chapters of a Norwegian electronic handbook for pharmaceutical interventions (Legemiddelhåndboka), which can assist practitioners in efficient and precise searching and enhance keywords based search interfaces.

The paper is structured as follows: Chapter 2 describes the architecture of our system. Chapter 3 explains methods used to solve the project tasks, while chapter 4 presents results. Chapter 5 provides a discussion of the results, while chapter 6 concludes the paper. Stopwords, medical terms and patient cases are listed in the Appendix.

# CHAPTER 2
# SYSTEM ARCHITECTURE

This chapter describes the architecture of our system. The system is build as a decision support system for practitioners at a hospital. To prevent any confusion, two roles are introduced.

**User** Practitioner that uses the system to search for relevant therapy chapters for a given patient case.

**System administrator** Computer savvy personnel, in this case us, that manages the system. Doing the parsing of input files and building indices.

Our system consist of three logical parts: Python modules, input files, and output files. There is also the external Python library Whoosh, used for indexing and searching. In the following sections, each part will be explained in detail. For an outline of the system, see Figure 2.1.

## 2.1 Whoosh

Whoosh[1] is a search-engine library written in Python to support fast indexing and searching of text collections. The library provides high performance, multifunctional queries and support for scoring algorithms.

We use it to store, index and search on the four types of input data we must handle: ICD-10 codes, ATC codes, therapy chapters, and patient cases. We only use Whoosh search functionality for the first two tasks.

[1] Whoosh: http://pypi.python.org/pypi/Whoosh

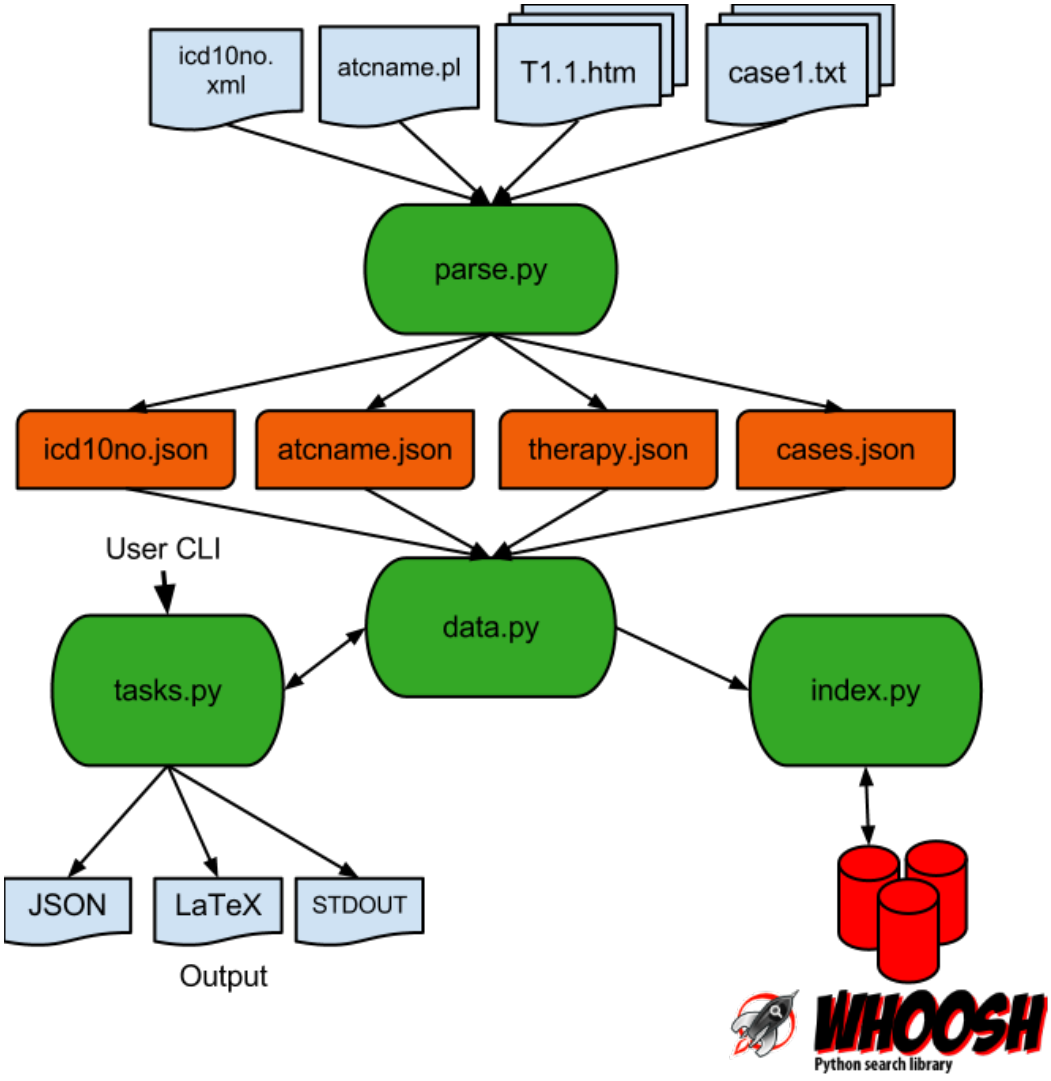


**Figure 2.1:** System overview

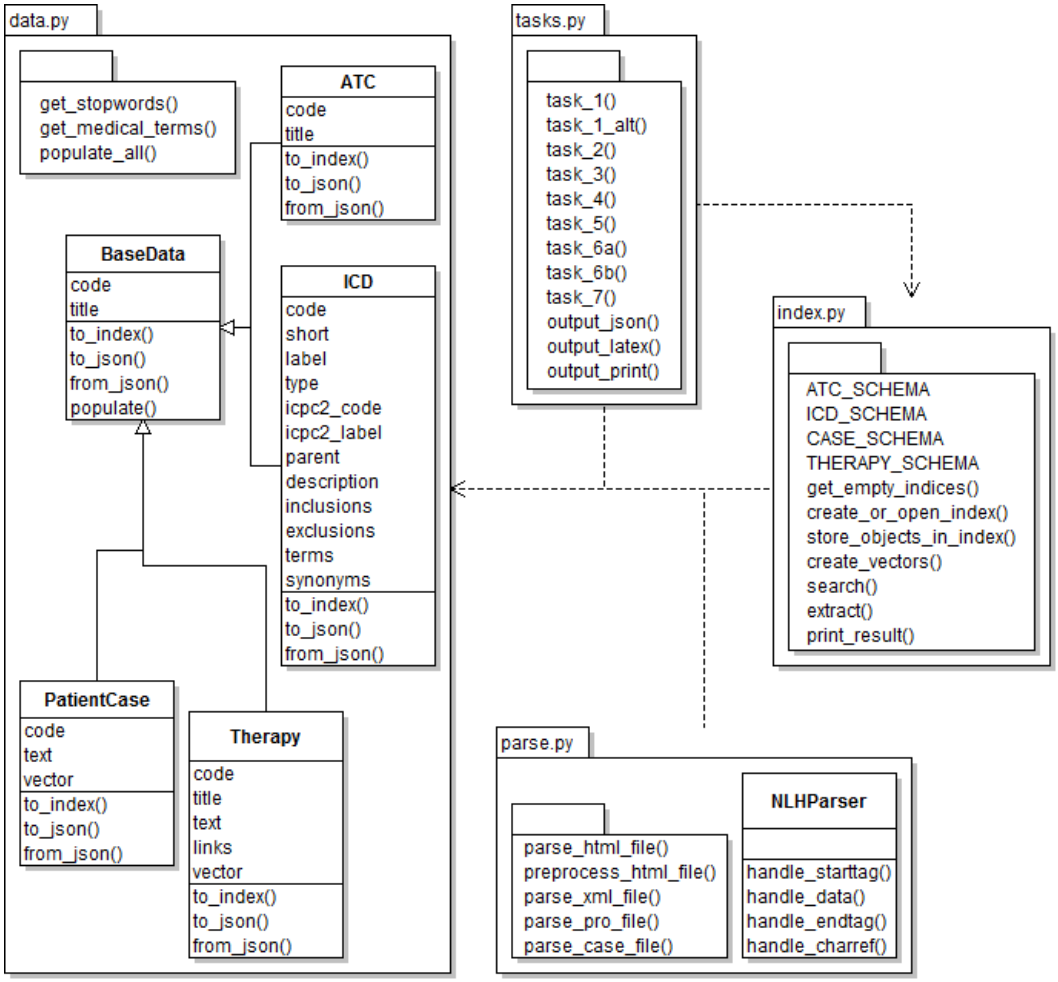


**Figure 2.2:** Class diagram of the Python modules

## 2.2 Python modules

The system consists of four Python modules: parse.py, data.py, index.py, and tasks.py. Each module has its own features and tasks, and together they form the functional core of the system. The external Python library, Whoosh, provides index and search functionality to the core. In the following paragraphs each module is described in greater detail. For a full overview of the modules with their associated classes, attributes and methods, see the system's class diagram in Figure 2.2.

**parse.py** This module preprocess and converts input files to the more preferred JSON format, which gives better readability and reduces complexity for the following modules, which now only have to support one type of file.

The module supports multiple file formats as input, for further information see the input files section 2.3. The module is especially important for stripping out all the HTML tags from the therapy chapters.

**data.py** The JSON files created by the parse module, are used as input to the data module. The data module holds representation of all the data the system needs. As a basis for holding the data we have the BaseData class, which is inherited by more specific classes for the different representations. As the system loads a JSON file, it determines which representation that should be used, ICD, ATC, PatientCase or Therapy.

**index.py** Main module for building and managing the indices. After the data module has created the representation and holds the data, the index module can build indexes of it. This gives the system the ability to store text and to search for terms.

**tasks.py** This module contains the Command Line Interface (CLI), which makes the user able to interact with the system. The module contains methods to perform and solve the different tasks of the assignment, specified by the user. Output is generated by this module, either as STDOUT print or as a JSON- or LaTeX-file. Sample output can be seen in the appendices.

## 2.3 Input files

The system needs to support multiple file formats to be able to preprocess and parse the files given in the assignment. The ICD-10 file is a .xml file, the ATC file is a .pro file, the "Norsk Legemiddel Håndbok" is HTML and the cases are .txt files. So the parse module has support for files ending with: .xml, .pro, .htm and .txt.

## 2.4 Output files

The result of the system is always presented in the command line. To be able to store the results and present them in this report, it was necessary and beneficial for us to make an output feature. The system is able to print the results to a JSON- or LaTeX-file. Storing results in JSON files are necessary to be able to use them for other tasks, such as task 6.

# CHAPTER 3

# METHOD

This section describes methods used to solve the tasks of the assignment, including preprocessing and parsing of input files.

The first two tasks are solved using a probabilistic method. The third task includes a more complex solution; both the probabilistic model and the relative proximity of terms are taken into account to find the best results. For task 6, we combine the results from prior tasks to achieve the most correct matching of documents for each patient case. This matching will be compared to a gold standard, to see how well our classification really is.

## 3.1 Preprocessing and parsing

We preprocess and parse input data files before we save them to disk as JSON files. This is to allow us to run a task that might be dependent on other tasks, without having to re-run previous tasks. We also store task results as JSON files.

### 3.1.1 ICD-10

ICD-10 were provided in an ownl2 file. We simply parsed it as a normal XML file, with the ElementTree module in the Python standard library. Nodes named umls_tui, umls_conceptId, and umls_atomId we throw away. Nodes named underterm, synonym, inclusion, and exclusion we store as list of strings. The rest of the nodes are handled as strings. All these values we store as attributes on a single ICD object. ICD-10 codes that lack both label and code_compacted we also throw away.

### 3.1.2 ATC

Parsing ATC is very straight forward. Each line represent a code, title pair where codes are not unique.

### 3.1.3 Therapy chapters

Therapy chapters from "Norsk legemiddelhåndbok" were provided as HTML files, invalid html5 files in iso-8859-1 charset. We first preprocessed these files by removing some of the HTML tags to make them easier to parse, and we converted them to utf-8 charset.

We created a custom parser for parsing therapy chapters, based on Python's HTMLParser. We parse one HTML file at a time, creating Therapy objects for each chapter or sub*-chapter. The text found in these chapters are stored on the objects. We stored links as a list on each object, while we preserved their text in the object text. Sections which list relevant drugs were removed from the text but stored as they might be useful later.

We manually removed sub-chapter T17.2 and T19.7 as they had no title nor contained any text.

### 3.1.4 Patient cases

Patient cases were provided as a Word file, which included eight cases. We created a text file for each case, and made sure they were in utf-8 charset.

## 3.2 Stopwords

To reduce the number of terms in documents, and to remove words which provide little or no relevant information value, we remove stopwords. We use a list of Norwegian stopwords in both Bokmål and Norwegian Nynorsk which we found online[1], and we added a few words ourself. A complete list of these stopwords can be found in Table A.1 in chapter A.

## 3.3 Task 1: Autocoding ICD-10

The first task was to find the most relevant ICD-10 code and therapy chapter for each sentence in the cases. The ICD-10 codes have multiple fields that can be used to make this classification. We decided to solve task 1a in two different ways.

**Method A** For each ICD-10 code, use only the label to find relevant codes for each sentence in the cases.

[1] Stopword source: http://www.wisweb.no/999/147/33899-170.html

**Method B** For each ICD-10 code, use all text to find relevant codes for each sentence in the cases.

We have defined all text as the following fields: label, terms, synonyms and inclusions.

**Scoring algorithm** To score and sort the results, we use the default scoring module given by Whoosh[2]. The module use the scoring algorithm BM25F to calculate the rankings of each code, and it returns a ranked list over relevant codes for each sentence.

BM25F[3] is a retrieval function that computes the best match for a given query over a set of documents. For each document in the set there is calculated a score based on inverse document frequency (IDF), term frequency(TF) and document length normalization. The terms are taken from an query $Q$, which can consist of several query terms $q_1, ..., q_n$. See Equation 3.1 for BM25F's mathematical equation.

$$\text{score}(D, Q) = \sum_{i=1}^{n} \text{IDF}(q_i) \cdot \frac{f(q_i, D) \cdot (k_1 + 1)}{f(q_i, D) + k_1 \cdot (1 - b + b \cdot \frac{|D|}{\text{avgdl}})} \tag{3.1}$$

$$\text{IDF}(q_i) = \log \frac{N - n(q_i) + 0.5}{n(q_i) + 0.5} \tag{3.2}$$

- Where $f(q_i, D)$ is $q_i's$ term frequency in the document $D$.
- $k_1$ and $b$ free parameters.
- $|D|$ the length of document $D$ in words.
- $avgdl$ the average document length.
- $N$ total number of documents in the set and $n(q_i)$ the number of documents containing $q_i$.

Applied to our task, the function uses the following as set of documents:

**1a.** The ICD-10 codes, each code representing a document.

**1b.** The therapy chapters, each chapter representing a document.

For each patient case, there are several sentences. A sentence forms one query, where the words in the sentence are classified as terms. By using BM25F, we get a ranked list of best matching documents to the sentence, and thereby an indication of what is wrong with the patient.

[2]Scoring module: http://packages.python.org/Whoosh/api/scoring.html#whoosh.scoring

[3]BM25F algorithm: http://en.wikipedia.org/wiki/Okapi_BM25

## 3.4 Task 2: Autocoding ATC

The second task is very similar to the first, the only difference is that the set of documents are ATC-classifications. As mentioned, each classification consist of a code and a title. The title is the field that we use to find the most relevant classification as described in task 1.

## 3.5 Task 3: Ranking using vector models

We decided to use a vector model for calculating the similarities between patient cases and therapy chapters. For each document, both patient cases and therapy chapters, we created a vector with all the terms and their TF-IDF value. These document-vectors gives us a pseudo term-document matrix, without having to create an actual matrix. As there were over 30,000 terms and almost 1000 documents, a full term-document matrix would have 30M cells. Our pseudo term-document only have 124,065 term-weight pairs.

We have tested four different varieties of TF-IDF. For TF we tested log normalization ($1 + \log f_{i,j}$) and raw frequency ($f_{i,j}$). For IDF we tested inverse frequency ($\log(\frac{N}{n_i})$) and probabilistic inverse frequency ($\log(\frac{N-n_i}{n_i})$).

**Method A** log normalization and inverse frequency as seen in Equation 3.3

**Method B** log normalization and probabilistic inverse frequency

**Method C** raw frequency and inverse frequency

**Method D** raw frequency and probabilistic inverse frequency

$f_{i,j}$ is the frequency of term $i$ in document $j$, $N$ is the total number of documents, while $n_i$ is the document frequency of term $i$.

$$w_{i,j} = \begin{cases} (1 + \log f_{i,j}) \times \log(\frac{N}{n_i}) & \text{if } f_{i,j} > 0 \\ 0 & \text{otherwise} \end{cases} \tag{3.3}$$

For each clinical note (which we have defined to be a patient case), we calculated the similarities with all therapy chapter vectors. The similarity is calculated as the *cosine of the angle* between the two vectors, as seen in Equation 3.4, where $d_j$ is the term vector of therapy chapter $j$, $q$ is a patient case vector and $t$ is the total number of terms.

$$\begin{aligned} sim(d_j, q) &= \frac{\vec{d_j} \bullet \vec{q}}{|\vec{d}| \times |\vec{q}|} \\ &= \frac{\sum_{i=1}^{t} w_{i,j} \times w_{i,q}}{\sqrt{\sum_{i=1}^{t} w_{i,j}^2} \times \sqrt{\sum_{i=1}^{t} w_{i,q}^2}} \end{aligned} \tag{3.4}$$

We sorted results based on their similarity score (highest first), and returned the first ten results.

## 3.6 Task 4: Evaluation

Evaluating results for a given query give great insight in what level of quality and correctness you can expect from an information retrieval system. This evaluation requires human relevance judgement. The problem is that it is a complex task, which is difficult and time consuming for humans to do. It would be beneficial to be able to automatically evaluate the results.

Automatic evaluation has the problem of deciding if a retrieved document is relevant or not without human intervention. As the queries in this task are patient cases, which is known beforehand, we can create a criteria for whether a document is relevant or not. We have simply taken all the words in the patient cases and created a list of those with medical importance, and if both the query and a retrieved document contains such a word we deem it relevant. These medical terms are listed in Table A.2 in Appendix.

We have calculated average precision at ten documents seen (P@10) and R-precision for the four different TF-IDF definitions A to D as described in section 3.5.

We have also calculated the Kendall tau coefficients for the four ranking methods. Kendall tau coefficient is a rank correlation metric for automatic comparing of two ranking methods to determine how differently one varies from another. It does not consider relevance of retrieved documents, only the relative ordering of two rankings.

## 3.7 Task 5: Exchange evaluations

Task 5 is to evaluate each task 4 method used by other groups. We have received reports of the results of task 4 from six other groups. None of the other groups seems to have actually implemented automatic evaluation of results. For each of the identified automatic evaluation methods, we have written a short evaluation.

**Group 2** suggest we compare results of task 3 with ICD and ATC matching, but task 3 match patient case with therapy chapters, while task 1 and 2 match patient case with ICD and ATC codes. While this might be a good idea, it is not obvious how they would achieve the matching by means of implementation.

**Group 3** propose we could calculate harmonic mean and E-measure, which would require calculating precision and recall. We are not sure how we would be able to calculate recall as we do not have knowledge of all relevant documents.

**Group 4** suggest we use user feedback for automatic evaluation of results.

**Group 5** suggest no method for automatic evaluation. They states that an automatic evaluation would be too subjective.

**Group 8** propose use of a reference corpus, which would require an expert in the field. They also suggest evaluation based on query expansion, for example relevance feedback with the Rocchio[4] algorithm.

**Group 14** suggest use of user feedback for evaluation, same as group 4.

## 3.8 Task 6: Improving the ranking

For part A we process task 1 and 2 results by giving each patient case a list of relevant ICD-10 and ATC codes, and giving each ICD-10 and ATC code a list of relevant therapy chapters. Each patient case can have the same code listed several times, one for each line potentially. Each code can also have the same chapter listed several times. Therefor we created a list of the relevant therapy chapters for each patient case based on task 1 and 2 results, with a weight for each with the number of times they are listed. We use the hierarchical structure of the codes to add relevant chapters for parent codes, but with a smaller weight (0.1) for each chapter.

We then used the hierarchical structure of therapy chapters to boost parent chapters if more than one if its sub-chapters were considered relevant. The parent chapter will get a weight of the max of its current weight and the weights if its children plus a bonus of the amount of children which are relevant.

For part B we simply merge the results of task 3 and task 6 A where we multiply the score of task 3 results by a constant of 200. The constant was selected to give both set of results fairly equal weight, so that chapters in both sets gets the highest score.

We will evaluate these results with the methods described in section 3.6, precision at ten documents seen and R-precision.

## 3.9 Task 7: Gold standard

We have been unable to complete this task as a gold standard has not been provided.

[4] Rocchio algorithm: http://en.wikipedia.org/wiki/Relevance_feedback

# CHAPTER 4
# RESULT

This chapter presents results of the preprocessing and parsing of input files and the results from the assignment tasks. For task 1 and 2 we only present a subset of the results.

## 4.1 Preprocessing and parsing

The results of parsing the different input files can be seen in Table 4.1. Each code, chapter and case is stored in an object, and saved to JSON files.

**Table 4.1:** Parsed object counts

| Type | Count |
|---|---|
| ATC codes | 7945 |
| ICD10 codes | 10521 |
| Patient cases | 8 |
| Therapy chapters | 917 |

Table 4.2 contains statistics showing the results of parsing "Norsk legemiddelhåndbok" HTML files — therapy chapters. The statistics lists the amount of chapters, in total and with text, for each of the different chapter types — from chapter to subsubsubsubchapter.

Parsing of patient cases are summarized in Table 4.3. Stopwords refers to stopwords in the case text which have been removed, terms is the number of unique terms (words) in the text.

We store input data and results in JSON format, so we can work with them without having to parse or produce them first. To demonstrate the effectiveness of this method we list and compare the time (in seconds) it takes to parse input data and loading JSON files in Table 4.4.

**Table 4.2:** Therapy chapters statistics

| Chapter type | Count | With text |
|---|---|---|
| Chapter | 24 | 24 |
| Subchapter | 153 | 104 |
| Subsubchapter | 384 | 336 |
| Subsubsubchapter | 329 | 326 |
| Subsubsubsubchapter | 27 | 27 |
| Total | 917 | 817 |

**Table 4.3:** Patient cases statistics

| Case # | Lines | Stopwords | Terms | Medical terms |
|---|---|---|---|---|
| 1 | 13 | 1 | 55 | 10 |
| 2 | 22 | 1 | 169 | 24 |
| 3 | 17 | 0 | 125 | 30 |
| 4 | 6 | 1 | 49 | 4 |
| 5 | 11 | 2 | 63 | 14 |
| 6 | 7 | 0 | 41 | 9 |
| 7 | 20 | 2 | 123 | 29 |
| 8 | 12 | 0 | 90 | 19 |
| Total | 108 | 7 | 715 | 125 |

**Table 4.4:** Comparing effectivenes of JSON

| Type | Parse time | JSON load time | Speedup |
|---|---|---|---|
| ATC codes | 0.209 | 0.101 | 52% |
| ICD10 codes | 9.496 | 0.549 | 94% |
| Patient cases | 0.008 | 0.001 | 88% |
| Therapy chapters | 11.326 | 0.177 | 98% |

## 4.2 Task 1: Autocoding ICD-10

A sentence can match zero to many ICD-10 codes, but only the most specific shall be claimed as a match and presented in the results. Whoosh gives us a ranked list of the codes, and from this ranking the most specific code is selected. If there is more than one ICD-10 code that scores high and the match scores of the top results are close, the top three results are presented. If there is no match, a . is printed in the result table. This also applies to the therapy chapters and ATC-classifications.

Autocoding of ICD-10 codes against patient case 1, 4, 5, and 6 can be seen in Table 4.5, which list relevant ICD-10 codes for each sentence. Table 4.6 list results for task 1 B (therapy chapter T1.1.1, T5.5, T8.9.2, and T24.2.1.7). Results for both method A and method B, described in section 3.3, are listed next to each other for easy comparison.

## 4.3 Task 2: Autocoding ATC

Autocoding of ATC codes against patient case 1, 2, 3, 7, and 8 can be seen in Table 4.7, where each sentence in patient cases are listed with relevant ATC codes. Table 4.8 list relevant ATC codes for therapy chapter T1.10, T2.2.5.1, T3.1, and T6.2.3.

## 4.4 Task 3: Ranking using vector models

Ranked lists of relevant therapy chapters for each patient case can be found in Table 4.9 and Table 4.10.

## 4.5 Task 4: Evaluation

An example of terms shared between retrieved therapy chapter and patient case can be seen in Table 4.11, where relevant medical terms are boldfaced. The patient case concerns a patient with diabetes mellitus and the first result "T3.1 Diabetes mellitus" is spot on, while the second result "T20.2.1. Bronkial astma" is not relevant at all. We have calculated average precision at ten documents seen (P@10) and R-precision, which are listed in Table 4.12.

Rank correlation metrics are an automatic way for comparing two ranking methods, to determine how differently one varies from another. It does not consider relevance of retrieved documents, only the relative ordering of two rankings. Table 4.13 list such a rank correlation metric, called Kendall tau coefficients.

**Table 4.5:** Task 1 A, patient case 1, 4, 5, and 6

| Clinical note | Sentence | Method A | Method B |
|---|---|---|---|
| 1 | 1 | E10-E14 | E12, E10-E14, E14 |
| | 2 | E10-E14, E23.2 | E14 |
| | 3 | Y61.3 | Y61.3, Y62.3, Y60.3 |
| | 4 | Q02, P92.3 | P92.3 |
| | 5 | Z97.2, Q26.3 | Z97.2 |
| | 6 | . | E10-E14 |
| | 7 | . | . |
| | 8 | . | . |
| | 9 | O36.6 | P03.5 |
| | 10 | . | L85.3 |
| | 11 | . | . |
| | 12 | Z01.3 | Z34 |
| | 13 | Z38.1, Z38.0, Z38.3 | Y87.2 |
| 4 | 1 | O96 | M08, O96, R06.2 |
| | 2 | O33 | Q38.4 |
| | 3 | P92.4 | Z58.6 |
| | 4 | R96.0, N88.1, H91.2 | H91.2, M23.2 |
| | 5 | I20 | I20.1, I20 |
| | 6 | O84.0 | O84.0 |
| 5 | 1 | O26.2, N88.1 | O26.2, N91.1, N91.4 |
| | 2 | R15, R19.5 | R19.5, R19.4 |
| | 3 | Y65.0 | C77.8 |
| | 4 | O46.9 | . |
| | 5 | D83 | D83 |
| | 6 | S63.1 | S60.0 |
| | 7 | I84.3 | H60.0 |
| | 8 | C86.6 | R85 |
| | 9 | Q56.4, Q56, F70 | Q56, D57.3, M93.9 |
| | 10 | D80.6 | D80.6 |
| | 11 | Y61.4, Y62.4, Y60.4 | Y61.4, Y62.4, Y60.4 |
| 6 | 1 | R98, R59 | R98, R59.9, R59.0 |
| | 2 | R76.2 | R76.2 |
| | 3 | . | B90.9 |
| | 4 | D83, U80.0 | U80 |
| | 5 | E59, E58, E60 | U80 |
| | 6 | Y84.4 | P24.3 |
| | 7 | G81.0, G82.3, G82.0 | G81.0, G82.3, G82.0 |

**Table 4.6:** Task 1 B, chapter T1.1.1, T5.5, T8.9.2, and T24.2.1.7

| Chapter | Sentence | Method A | Method B |
|---|---|---|---|
| T1.1.1 | 1 | B01, B01.9, B01.8 | B01.9, B01, B01.8 |
| | 2 | Z20, Z20.8, Z20.9 | A88.0, Z20, Z20.0 |
| | 3 | B01, B01.9, B01.8 | B01.9, B02.9, B02.8 |
| | 4 | C21.2 | O69.8 |
| | 5 | Q90.2 | C92.5 |
| | 6 | N92.4 | Z00.2 |
| | 7 | R00-R99 | R00-R99 |
| | 8 | P36.2 | A41.0, B95.6, G11.9 |
| | 9 | G11.1 | B01.2†, G11.1, G11.9 |
| | 10 | B01 | B01, B01.9, B01.8 |
| T5.5 | 1 | F68.0, Z53.8, Z41.9 | R19.6 |
| | 2 | U00-U49, Z31.5, Q95.4 | Z90.5, U00-U49, I25.2 |
| | 3 | F32.8 | F33, F32.8 |
| | 4 | F51.2 | R41.8, R41, E23 |
| | 5 | F41.0, F32.3, F32.2 | F32.2, F32.3, F33.2 |
| | 6 | F31.3 | F33.3, F33.2 |
| | 7 | F31.8 | R70.0, F52, I51 |
| | 8 | F31.4, F31.5 | F33.3 |
| | 9 | Z34.9, Q97.1, F32.8 | Z34.9, O84.0 |
| | 10 | Z29, Z29.9 | F33.3, F31.8, Z29 |
| | 11 | O15.9 | Y4N |
| | 12 | Z55.0 | R45.4, R19.6, R46.0 |
| | 13 | P91.4, F20.4 | F60-F69 |
| T8.9.2 | 1 | O42.1, O42.0 | O02.9, O02.8, G45.9 |
| | 2 | O96, Z00.8, I63 | O96, I63.3 |
| | 3 | G45.9, J46 | A50.0 |
| | 4 | Z00, A52.2, I48 | Z00, Z92.2, Z03 |
| | 5 | B20-B24 | I67, I68.0 *, I67.8 |
| T24.2.1.7 | 1 | Z56.6 | F43 |
| | 2 | K59.1, O34, O35 | I22 |
| | 3 | I20 | I20.0, I20.1, I22 |
| | 4 | P08.0, Y61.3, Y62.3 | Y61.3, Y62.3, Y60.3 |
| | 5 | O63, P92.5, L21.1 | O63, R68.1, U00-U99 |
| | 6 | J46, I20, I20.0 | I20.0, I44.1 |
| | 7 | O42.1, O42.0 | F51.2, R96.1, F20.6 |
| | 8 | O42.1, O42.0, Z39.0 | T80.1, T80.2, T80 |
| | 9 | T32.3 | Z53 |

**Table 4.7:** Task 2 A, patient case 1, 2, 3, 7, and 8

| Case | Sentence | ATC codes | Sentence | ATC codes |
|---|---|---|---|---|
| 1 | 1 | A10X | 8 | . |
| | 2 | A10X | 9 | A10AD |
| | 3 | A10AE | 10 | . |
| | 4 | . | 11 | . |
| | 5 | A10AB1 | 12 | . |
| | 6 | . | 13 | . |
| | 7 | . | | |
| 2 | 1 | . | 12 | . |
| | 2 | . | 13 | . |
| | 3 | . | 14 | A10AD1 |
| | 4 | . | 15 | . |
| | 5 | N1BB1 | 16 | . |
| | 6 | . | 17 | . |
| | 7 | D8AC2 | 18 | . |
| | 8 | . | 19 | . |
| | 9 | . | 20 | V3AN5, G3AA7, G3AA12 |
| | 10 | . | 21 | R3BB1, R3BA2, R3AC3 |
| | 11 | V4CX | 22 | A10AD1 |
| 3 | 1 | C1BA | 10 | . |
| | 2 | . | 11 | . |
| | 3 | . | 12 | J1CE1, D6AX2, D10AF3 |
| | 4 | J7BB1 | 13 | N5BA1 |
| | 5 | . | 14 | . |
| | 6 | A10AD1 | 15 | . |
| | 7 | . | 16 | . |
| | 8 | . | 17 | . |
| | 9 | . | | |
| 7 | 1 | . | 11 | N2 |
| | 2 | V9B | 12 | N2BE1, M1AE2, M1AE2 |
| | 3 | . | 13 | M1AE2, M1AE2 |
| | 4 | . | 14 | N2AA59, N2AA59 |
| | 5 | . | 15 | N2AB1 |
| | 6 | . | 16 | B5BB |
| | 7 | . | 17 | V10B |
| | 8 | . | 18 | N2AA1 |
| | 9 | . | 19 | N2AA1, N2AA1 |
| | 10 | . | 20 | N2A, Z9OP, Z9SA |
| 8 | 1 | C1BA | 7 | V4CB |
| | 2 | . | 8 | V4CB |
| | 3 | A7AA, D1AA, G1AA | 9 | . |
| | 4 | V4CB | 10 | A12AA12 |
| | 5 | . | 11 | R1AX |
| | 6 | J1CE2, J1CE1 | 12 | C2N |

**Table 4.8:** Task 2 B, chapter T1.10, T2.2.5.1, T3.1, and T6.2.3

| Chapter | Sentence | ATC codes | Sentence | ATC codes |
|---|---|---|---|---|
| T1.10 | 1 | C5A, D5A, A1AB | 6 | J1EE1, J1RA1 |
| | 2 | V7, V7A | 7 | C10AC, N5BA1, M4AB |
| | 3 | V9D, J1EB, A7EC1 | 8 | A5AB, D5BB, D10AF |
| | 4 | J7BC20, A7EC1, N4AC | 9 | V4CB, V4CD, V4CG |
| | 5 | B1AD12, B5A, C8D | 10 | J7BC20, A7EC1 |
| T2.2.5.1 | 1 | M4AA, L3AB1, L3AB4 | 2 | A5AB, D5BB, D10AF |
| T3.1 | 1 | C2LG51, A10X, V3AA | 23 | V3AH |
| | 2 | C10AC, M4AB, H2AB | 24 | H4AA1, Z0CA, A5AB |
| | 3 | B5D, B5BA3, B5CX1 | 25 | A5AB, D5BB, D10AF |
| | 4 | A10BX2, A10BX3, Z0ET | 26 | B5BA3, B5CX1, V4CA2 |
| | 5 | B5BB, G3, M5B | 27 | V9DX1 |
| | 6 | V6DB | 28 | A14A |
| | 7 | A10AD1 | 29 | B5BC2, D2AE1, B5BA3 |
| | 8 | A10BA2, Z9MF, A10BD2 | 30 | A10AE, B5BB3, B5BB |
| | 9 | A10AE | 31 | V7AD |
| | 10 | A10BD, J7BC20, A10AD1 | 32 | A10AB1, A10AD4 |
| | 11 | V3AH | 33 | . |
| | 12 | A7EC1, N4AC | 34 | V3AH |
| | 13 | B5BA3, B5CX1, V4CA2 | 35 | A10X, N4AC, N7 |
| | 14 | A10AE | 36 | C5BA, C5AX, C5BB |
| | 15 | A10BD3 | 37 | Z9AC |
| | 16 | A10AE4, A10AE5 | 38 | Z9ST, V9G |
| | 17 | A10AE | 39 | A5AX, M4AC, A7EC1 |
| | 18 | A10AC1, N4AC | 40 | Z9A2 |
| | 19 | C9AA5, A10AB, A10B | 41 | C10AC, M4AB, A10BA2 |
| | 20 | A10AC1, A10AB | 42 | . |
| | 21 | J4AK, S1KX, P1A | 43 | Z9AC |
| | 22 | V3AH, J4AK, M9AX | 44 | J4AK, V4CB, V4CD |
| | | | 45 | A1AB |
| T6.2.3 | 1 | V3AA, N4AC | 4 | V9G, V4CB, V4CD |
| | 2 | S3, V4CJ, A7EC1 | 5 | V3AA, N3AX12, Z9BD |
| | 3 | N4AC | | |

**Table 4.9:** Task 3 results (part 1)

| Case | Rank | Score | Relevant chapter |
|---|---|---|---|
| 1 | 1 | 0.08 | T3.1: Diabetes mellitus |
| | 2 | 0.04 | T10.2.1: Bronkial astma |
| | 3 | 0.04 | T14.5.1: Polycystisk ovarialt syndrom (PCOS) |
| | 4 | 0.04 | T23.1.1.2: Faste og stress |
| | 5 | 0.04 | T5.4.1: Schizofreni |
| | 6 | 0.03 | T14.2.1: Forskyvning av normal menstruasjon |
| | 7 | 0.03 | T10.2.1.1: Mild og moderat astma |
| | 8 | 0.03 | T18.1.4: Kontroll og oppfølging |
| | 9 | 0.03 | T16.13.1: Generalisert kløe |
| | 10 | 0.03 | T9.1.5: Anafylaktoide reaksjoner |
| 2 | 1 | 0.08 | T10.2: Obstruktiv lungesykdom |
| | 2 | 0.06 | T10.2.2: Kronisk obstruktiv lungesykdom (kols) |
| | 3 | 0.05 | T8.4.1.2.2: Atrioventrikulær nodal reentrytakykardi |
| | 4 | 0.05 | T10.2.1: Bronkial astma |
| | 5 | 0.04 | T8.3.2.2: Hjerteinfarkt med ST-elevasjon |
| | 6 | 0.04 | T3.1: Diabetes mellitus |
| | 7 | 0.04 | T10.8: Sarkoidose |
| | 8 | 0.04 | T15.3.7: Liten melkeproduksjon |
| | 9 | 0.03 | T5.3.1.3: Alkohol abstinensreaksjoner |
| | 10 | 0.03 | T6.2.2: Klasehodepine («Cluster headache») |
| 3 | 1 | 0.09 | T1.10: Akutt bakteriell meningitt |
| | 2 | 0.05 | T3.1: Diabetes mellitus |
| | 3 | 0.05 | T8.1: Hypertensjon |
| | 4 | 0.05 | T1.11: Bakteriell endokarditt |
| | 5 | 0.04 | T16.7.1: Skabb |
| | 6 | 0.04 | T8.2.1: Malign hypertensjon |
| | 7 | 0.04 | T19.1: Feber |
| | 8 | 0.04 | T8.2.2: Hypertensjonsencefalopati |
| | 9 | 0.04 | T8.3.2.2: Hjerteinfarkt med ST-elevasjon |
| | 10 | 0.04 | T14.6.4: Akutt bekkeninfeksjon |
| 4 | 1 | 0.09 | T8.3: Koronarsykdom |
| | 2 | 0.06 | T11.1.1.4.7: Emosjonell rhinitt |
| | 3 | 0.05 | T8.2.4: Hypertensjonskrise og hjerteinfarkt eller ustabil angina |
| | 4 | 0.05 | T4.6.3: Arteriell trombose |
| | 5 | 0.04 | T8.3.1: Stabil koronarsykdom (stabil angina pectoris) |
| | 6 | 0.04 | T8.4.1.2: Paroksystisk supraventrikulær takykardi |
| | 7 | 0.04 | T10.2.2: Kronisk obstruktiv lungesykdom (kols) |
| | 8 | 0.04 | T8.3.2.1: Ustabil angina/hjerteinfarkt uten ST-elevasjon |
| | 9 | 0.04 | T24.2.1.7: Myokardscintigrafi |
| | 10 | 0.03 | T15.3.7: Liten melkeproduksjon |

**Table 4.10:** Task 3 results (part 2)

| Case | Rank | Score | Relevant chapter |
|---|---|---|---|
| 5 | 1 | 0.06 | T12.10.1: Hemoroider |
| | 2 | 0.06 | T12.9.3: Dyschezi (rektumobstipasjon) |
| | 3 | 0.05 | T4.1: Anemier |
| | 4 | 0.05 | T1.6.2.1: Clostridium difficile enterokolitt |
| | 5 | 0.05 | T12.10.3: Fissura ani |
| | 6 | 0.05 | T12.11: Familiær adenomatøs polypose |
| | 7 | 0.05 | T5.5: Depresjoner |
| | 8 | 0.04 | T15.1.5: Svangerskapsindusert hypertensjon |
| | 9 | 0.04 | T4.1.3.2: Talassemi |
| | 10 | 0.04 | T13.2.5: Nevrogene blæreforstyrrelser |
| 6 | 1 | 0.06 | T2.2.5.1: Cancer i nyreparenkym og binyre |
| | 2 | 0.05 | T11.3.2.2: Kronisk tonsillitt |
| | 3 | 0.04 | T11.3.1.2: Kronisk faryngitt |
| | 4 | 0.04 | T1.7.7: Lymfogranuloma venereum |
| | 5 | 0.04 | T11.4.4: Halitosis |
| | 6 | 0.04 | T1.1.8: Skarlagensfeber |
| | 7 | 0.03 | T11.3.2.1: Akutt tonsillitt |
| | 8 | 0.03 | T1.6.1: Ikke-inflammatoriske, toksinpregete enteritter |
| | 9 | 0.03 | T10.3.4: Pneumonier, bakterielle og med ukjent etiologi |
| | 10 | 0.03 | T10.2.1.1: Mild og moderat astma |
| 7 | 1 | 0.08 | T6.2.3: Spenningshodepine (Tensjonshodepine) |
| | 2 | 0.07 | T20.2.1: Akutte smerter |
| | 3 | 0.06 | T21.1.1.2: Nevropatiske smerter |
| | 4 | 0.06 | T20.2.3.1: Praktisk gjennomføring av smertebehandling hos pasienter med kort livsprognose |
| | 5 | 0.06 | T22.4.1.1: Postoperativ grunnanalgesi |
| | 6 | 0.06 | T20.1.2.2: Opioidanalgetika |
| | 7 | 0.06 | T20.2.2.1: Praktisk gjennomføring av smertebehandling hos pasienter med antatt normal levetid |
| | 8 | 0.06 | T21.1.1.1: Nociseptive smerter |
| | 9 | 0.05 | T20.2.3.2: Bruk av sterkere opioider hos pasienter med kort livsprognose |
| | 10 | 0.04 | T6.5.1: Multippel sklerose |
| 8 | 1 | 0.08 | T11.3.2.1: Akutt tonsillitt |
| | 2 | 0.06 | T1.1.8: Skarlagensfeber |
| | 3 | 0.04 | T1.7.5: Syfilis |
| | 4 | 0.04 | T11.3.1.1: Akutt faryngitt |
| | 5 | 0.04 | T1.3: Mononukleose |
| | 6 | 0.04 | T1.10: Akutt bakteriell meningitt |
| | 7 | 0.03 | T16.5.1: Pyodermier |
| | 8 | 0.03 | T11.1.2.1: Akutt rhinosinusitt |
| | 9 | 0.03 | T10.3.4: Pneumonier, bakterielle og med ukjent etiologi |
| | 10 | 0.03 | T1.11: Bakteriell endokarditt |

**Table 4.11:** Task 4 shared terms (patient case 1)

| Rank | Chapter | Score | Relevant | Terms |
|---|---|---|---|---|
| 1 | T3.1 | 0.0832 | Yes | bruker, delvis, henvisning, hatt, **aceton-lukt**, injeksjon, år, hurtigvirkende, hvert, **mellitus**, siste, lite, hurtig, **insulin**, håndtere, synes, flere, dessuten, vurderer, **diabetes**, kvelden, måltid, langtidsvirkende, **blodtrykk**, normalt, døgn, sykehus |
| 2 | T10.2.1 | 0.0429 | No | bruker, delvis, hurtig, flere, dessuten, vurderer |
| 3 | T14.5.1 | 0.0407 | Yes | **insulin**, uteblir |
| 4 | T23.1.1.2 | 0.0372 | Yes | lite, **insulin**, dessuten, normalt, døgn |
| 5 | T5.4.1 | 0.0372 | Yes | delvis, år, fått, hvert, lite, håndtere, synes, flere, **diabetes** |
| 6 | T14.2.1 | 0.0332 | No | bruker, siste, brukt, flere, tatt |
| 7 | T10.2.1.1 | 0.0325 | No | bruker, delvis, henvisning, hatt, år, fått, hvert, hurtig, synes, flere, langtidsvirkende, døgn |
| 8 | T18.1.4 | 0.0309 | No | hvert, kontroller |
| 9 | T16.13.1 | 0.0304 | Yes | tørr, huden, **mellitus**, flere, **diabetes** |
| 10 | T9.1.5 | 0.0290 | Yes | **injiserer**, hatt, injeksjon, år, hurtig, **blodtrykk**, sykehus |

**Table 4.12:** Task 4 precision of each patient case search

| | Precision @ 10 | | | | R-precision | | | |
|---|---|---|---|---|---|---|---|---|
| Case | A | B | **C** | D | A | B | **C** | **D** |
| 1 | 60% | 60% | 70% | 50% | 0.67 | 0.67 | 0.71 | 0.60 |
| 2 | 50% | 60% | 70% | 60% | 0.80 | 0.67 | 0.71 | 0.83 |
| 3 | 90% | 80% | 80% | 80% | 0.89 | 0.88 | 0.75 | 0.75 |
| 4 | 70% | 70% | 60% | 70% | 0.71 | 0.86 | 0.83 | 0.86 |
| 5 | 80% | 80% | 90% | 90% | 0.88 | 0.88 | 0.89 | 0.89 |
| 6 | 80% | 80% | 80% | 80% | 0.75 | 0.75 | 0.88 | 0.88 |
| 7 | 100% | 100% | 100% | 100% | 1.00 | 1.00 | 1.00 | 1.00 |
| 8 | 90% | 90% | 80% | 80% | 0.89 | 0.89 | 0.88 | 0.88 |
| Avg | 77.5% | 77.5% | **78.8%** | 76.2% | 0.82 | 0.82 | **0.83** | **0.83** |

**Table 4.13:** Task 4 Kendall tau coefficients

| Case | A vs B | A vs C | A vs D | B vs C | B vs D | C vs D |
|---|---|---|---|---|---|---|
| 1 | 0.981 | 0.941 | 0.947 | 0.931 | 0.944 | 0.978 |
| 2 | 0.984 | 0.937 | 0.940 | 0.931 | 0.939 | 0.981 |
| 3 | 0.989 | 0.948 | 0.950 | 0.945 | 0.950 | 0.988 |
| 4 | 0.975 | 0.957 | 0.959 | 0.942 | 0.962 | 0.971 |
| 5 | 0.985 | 0.951 | 0.952 | 0.945 | 0.953 | 0.983 |
| 6 | 0.971 | 0.954 | 0.954 | 0.937 | 0.960 | 0.964 |
| 7 | 0.981 | 0.944 | 0.944 | 0.936 | 0.946 | 0.978 |
| 8 | 0.985 | 0.943 | 0.948 | 0.935 | 0.944 | 0.984 |
| Avg | 0.981 | 0.947 | 0.949 | 0.938 | 0.950 | 0.979 |

## 4.6 Task 5: Exchange evaluations

We have calculated precision at ten documents retrieved and R-precision for the groups which published their results of task 3, including our own results. Precision at ten can be seen in Table 4.14 while Table 4.15 list R-precision. Group 2 was not included as their results did not contain any therapy-chapters. It is important to point out that other groups might have handled parsing of therapy-chapters differently, which would affect greatly the content of chapters and therefor the rankings. For example the content of a sub-chapter could be included or excluded in the parent chapter, we use the latter option.

**Table 4.14:** Task 5 precision at ten documents retrieved

| Case | Group 1 | Group 3 | Group 4 | Group 5 | Group 14 |
|---|---|---|---|---|---|
| 1 | 60% | 60% | 40% | 80% | 60% |
| 2 | 50% | 70% | 50% | 80% | 40% |
| 3 | 90% | 60% | 70% | 90% | 30% |
| 4 | 70% | 30% | 50% | 70% | 20% |
| 5 | 80% | 60% | 50% | 70% | 80% |
| 6 | 80% | 30% | 50% | 80% | 10% |
| 7 | 100% | 40% | 30% | 80% | 50% |
| 8 | 90% | 50% | 70% | 80% | 50% |
| Avg | 77.5% | 50.0% | 51.2% | 78.8% | 42.5% |

## 4.7 Task 6: Improving the ranking

Results of ranking relevant therapy chapters with only using task 1 and 2 results (task 6 A) can be seen in Table 4.16 and Table 4.17. These results merged with task 3 results (task 6 B) is listed in Table 4.18 and Table 4.19.

We have calculated precision at ten documents seen and R-precision for

**Table 4.15:** Task 5 R-precision

| Case | Group 1 | Group 3 | Group 4 | Group 5 | Group 14 |
|---|---|---|---|---|---|
| 1 | 0.67 | 0.67 | 0.50 | 0.75 | 0.67 |
| 2 | 0.80 | 0.71 | 0.60 | 0.75 | 0.25 |
| 3 | 0.89 | 0.67 | 0.71 | 0.89 | 0.33 |
| 4 | 0.71 | 1.00 | 0.60 | 0.71 | 0.00 |
| 5 | 0.88 | 0.67 | 0.40 | 0.71 | 0.75 |
| 6 | 0.75 | 0.00 | 0.80 | 0.75 | 0.00 |
| 7 | 1.00 | 0.50 | 0.00 | 0.88 | 0.20 |
| 8 | 0.89 | 0.60 | 0.86 | 0.88 | 0.20 |
| Avg | 0.82 | 0.60 | 0.56 | 0.79 | 0.30 |

both task 6 A and B, which are listed in Table 4.20. Table 4.21 lists Kendall tau coefficients between the results of the three ranking methods.

**Table 4.16:** Task 6 A results (part 1)

| Case | Rank | Score | Relevant chapter |
|---|---|---|---|
| 1 | 1 | 22.90 | T3: Endokrine sykdommer |
| | 2 | 22.20 | T3.1: Diabetes mellitus |
| | 3 | 13.10 | T24.2: Nukleærmedisin |
| | 4 | 12.40 | T24.2.1: Nukleærmedisinsk diagnostikk |
| | 5 | 10.40 | T24.2.1.10: Nyrescintigrafi |
| | 6 | 10.40 | T24.2.1.13: Skjelettscintigrafi |
| | 7 | 8.20 | T3.2.1: Hypersekresjonstilstander |
| | 8 | 7.90 | T12: Mage-tarmsykdommer |
| | 9 | 7.80 | T24.2.1.19: Okkult tumor |
| | 10 | 7.50 | T3.2.1.3: Hypofysært betinget Cushings syndrom |
| 2 | 1 | 13.30 | T3.1: Diabetes mellitus |
| | 2 | 9.80 | T10: Nedre luftveissykdommer |
| | 3 | 9.70 | T1: Infeksjonssykdommer |
| | 4 | 8.80 | T15: Graviditet, fødsel og amming |
| | 5 | 8.80 | T24.2: Nukleærmedisin |
| | 6 | 8.50 | T10.2: Obstruktiv lungesykdom |
| | 7 | 8.30 | T17.1: Betennelsesaktige, revmatiske sykdommer |
| | 8 | 8.30 | T11: Sykdommer i øvre luftveier, øre, munn og svelg |
| | 9 | 8.20 | T14.1.1.1: Livmorinnlegg |
| | 10 | 8.10 | T15.3: Amming |
| 3 | 1 | 9.80 | T1: Infeksjonssykdommer |
| | 2 | 9.60 | T6: Nevrologiske sykdommer |
| | 3 | 8.90 | T6.1: Epilepsi, feberkramper |
| | 4 | 8.30 | T1.2: Influensa |
| | 5 | 8.20 | T6.1.2: Feberkramper |
| | 6 | 6.30 | T17: Muskel- og skjelettsykdommer |
| | 7 | 6.10 | T3.1: Diabetes mellitus |
| | 8 | 5.80 | T24.2: Nukleærmedisin |
| | 9 | 5.60 | T17.1: Betennelsesaktige, revmatiske sykdommer |
| | 10 | 5.10 | T24.2.1: Nukleærmedisinsk diagnostikk |
| 4 | 1 | 16.40 | T8: Hjerte- og karsykdommer |
| | 2 | 15.30 | T8.3: Koronarsykdom |
| | 3 | 14.60 | T8.3.1: Stabil koronarsykdom (stabil angina pectoris) |
| | 4 | 11.30 | T8.3.2: Ustabil koronarsykdom (ustabil angina) |
| | 5 | 10.60 | T8.3.2.2: Hjerteinfarkt med ST-elevasjon |
| | 6 | 6.90 | T24.2: Nukleærmedisin |
| | 7 | 6.20 | T24.2.1: Nukleærmedisinsk diagnostikk |
| | 8 | 5.70 | T3.1: Diabetes mellitus |
| | 9 | 5.60 | T1: Infeksjonssykdommer |
| | 10 | 5.30 | T17.1: Betennelsesaktige, revmatiske sykdommer |

**Table 4.17:** Task 6 A results (part 2)

| Case | Rank | Score | Relevant chapter |
|---|---|---|---|
| 5 | 1 | 10.50 | T1: Infeksjonssykdommer |
| | 2 | 9.00 | T1.6: Infeksiøse enteritter |
| | 3 | 8.80 | T24.2: Nukleærmedisin |
| | 4 | 8.20 | T1.6.2: Bakterielle, inflammatoriske enteritter |
| | 5 | 8.10 | T24.2.1: Nukleærmedisinsk diagnostikk |
| | 6 | 8.10 | T12: Mage-tarmsykdommer |
| | 7 | 7.20 | T1.6.2.4: Shigellose |
| | 8 | 7.00 | T23.3.1.1: Hypovolemisk sjokk |
| | 9 | 6.60 | T12.10: Anorektale forstyrrelser |
| | 10 | 6.50 | T6: Nevrologiske sykdommer |
| 6 | 1 | 4.60 | T7.9: Øyeskader |
| | 2 | 4.10 | T1: Infeksjonssykdommer |
| | 3 | 3.90 | T7.9.2: Perforerende skader (øye) |
| | 4 | 3.60 | T10: Nedre luftveissykdommer |
| | 5 | 3.50 | T11.4: Tenner, munnsykdommer og plager |
| | 6 | 2.90 | T1.7: Seksuelt overførbare infeksjoner (Soi) |
| | 7 | 2.80 | T10.3: Akutte infeksjoner i nedre luftveier og lunger |
| | 8 | 2.70 | T11.4.7: Akutt nekrotiserende gingivitt |
| | 9 | 2.40 | T4.4.1: Defekt blodplatefunksjon |
| | 10 | 2.40 | T16.9: Kutane bivirkninger av systemiske legemidler |
| 7 | 1 | 15.30 | T8.3.2.2: Hjerteinfarkt med ST-elevasjon |
| | 2 | 11.80 | T21.1: Lindring av smerter og andre plager i palliativ |
| | 3 | 11.30 | T22.4: Postoperativ fase |
| | 4 | 11.10 | T21.1.1: Smerter |
| | 5 | 10.50 | T22.4.1: Postoperativ smertebehandling |
| | 6 | 10.30 | T21.1.1.1: Nociseptive smerter |
| | 7 | 9.70 | T22.4.1.3: Opioider i postoperativ smertebehandling |
| | 8 | 9.10 | T20: Smerter |
| | 9 | 8.50 | T15: Graviditet, fødsel og amming |
| | 10 | 8.40 | T20.2: Akutte og kroniske smerter |
| 8 | 1 | 11.30 | T1: Infeksjonssykdommer |
| | 2 | 9.60 | T1.5: Urinveisinfeksjoner |
| | 3 | 9.30 | T24.2: Nukleærmedisin |
| | 4 | 8.90 | T1.5.1: Nedre urinveisinfeksjon |
| | 5 | 8.60 | T24.2.1: Nukleærmedisinsk diagnostikk |
| | 6 | 7.30 | T12: Mage-tarmsykdommer |
| | 7 | 7.00 | T24.2.1.16: Okkult bakteriell infeksjon. Inflammatorisk tarm |
| | 8 | 6.90 | T15: Graviditet, fødsel og amming |
| | 9 | 6.60 | T12.5: Galleveissykdommer |
| | 10 | 6.20 | T15.3: Amming |

**Table 4.18:** Task 6 B results (part 1)

| Case | Rank | Score | Relevant chapter |
|---|---|---|---|
| 1 | 1 | 38.20 | T3.1: Diabetes mellitus |
| | 2 | 22.90 | T3: Endokrine sykdommer |
| | 3 | 17.10 | T24.2: Nukleærmedisin |
| | 4 | 13.50 | T12.4.2: Hemokromatose |
| | 5 | 12.40 | T24.2.1.10: Nyrescintigrafi |
| | 6 | 12.40 | T24.2.1.13: Skjelettscintigrafi |
| | 7 | 12.40 | T24.2.1: Nukleærmedisinsk diagnostikk |
| | 8 | 12.30 | T12.2.2: Kronisk pankreatitt |
| | 9 | 11.50 | T3.2.1.3: Hypofysært betinget Cushings syndrom |
| | 10 | 11.30 | T16.13.1: Generalisert kløe |
| 2 | 1 | 24.50 | T10.2: Obstruktiv lungesykdom |
| | 2 | 21.30 | T3.1: Diabetes mellitus |
| | 3 | 19.80 | T10.2.2: Kronisk obstruktiv lungesykdom (kols) |
| | 4 | 15.00 | T15.3.7: Liten melkeproduksjon |
| | 5 | 14.20 | T14.1.1.1: Livmorinnlegg |
| | 6 | 13.70 | T10.2.1: Bronkial astma |
| | 7 | 13.40 | T8.4.1.1.1: Kronisk atrieflimmer |
| | 8 | 11.40 | T8.4.1.2.2: Atrioventrikulær nodal reentrytakykardi |
| | 9 | 11.30 | T24.2.1.7: Myokardscintigrafi |
| | 10 | 11.30 | T24.2.1.2: Dopamin transporter ligand scintigrafi |
| 3 | 1 | 21.00 | T1.10: Akutt bakteriell meningitt |
| | 2 | 16.30 | T1.2: Influensa |
| | 3 | 16.20 | T6.1.2: Feberkramper |
| | 4 | 16.10 | T3.1: Diabetes mellitus |
| | 5 | 12.90 | T7.8.2: Glaukom med åpen kammervinkel |
| | 6 | 10.80 | T15.1.4: Kronisk hypertensjon og svangerskap |
| | 7 | 10.80 | T1.7.5: Syfilis |
| | 8 | 10.30 | T8.1: Hypertensjon |
| | 9 | 10.20 | T1.11: Bakteriell endokarditt |
| | 10 | 9.90 | T8.3.2.2: Hjerteinfarkt med ST-elevasjon |
| 4 | 1 | 33.30 | T8.3: Koronarsykdom |
| | 2 | 22.60 | T8.3.1: Stabil koronarsykdom (stabil angina pectoris) |
| | 3 | 18.40 | T8: Hjerte- og karsykdommer |
| | 4 | 17.30 | T8.3.2: Ustabil koronarsykdom (ustabil angina) |
| | 5 | 16.60 | T8.3.2.2: Hjerteinfarkt med ST-elevasjon |
| | 6 | 12.30 | T24.2.1.7: Myokardscintigrafi |
| | 7 | 12.00 | T11.1.1.4.7: Emosjonell rhinitt |
| | 8 | 11.70 | T3.1: Diabetes mellitus |
| | 9 | 11.20 | T8.2.4: Hypertensjonskrise og hjerteinfarkt |
| | 10 | 10.10 | T4.6.3: Arteriell trombose |

**Table 4.19:** Task 6 B results (part 2)

| Case | Rank | Score | Relevant chapter |
|---|---|---|---|
| 5 | 1 | 17.80 | T12.10.1: Hemoroider |
| | 2 | 16.00 | T1.6.2.1: Clostridium difficile enterokolitt |
| | 3 | 14.60 | T4.1: Anemier |
| | 4 | 13.60 | T12.11: Familiær adenomatøs polypose |
| | 5 | 13.60 | T12.10.3: Fissura ani |
| | 6 | 13.20 | T1.6.2.4: Shigellose |
| | 7 | 12.40 | T12.9.3: Dyschezi (rektumobstipasjon) |
| | 8 | 12.10 | T24.2.1: Nukleærmedisinsk diagnostikk |
| | 9 | 11.50 | T5.5: Depresjoner |
| | 10 | 11.00 | T23.3.1.1: Hypovolemisk sjokk |
| 6 | 1 | 12.40 | T11.3.2.2: Kronisk tonsillitt |
| | 2 | 12.00 | T2.2.5.1: Cancer i nyreparenkym og binyre |
| | 3 | 9.00 | T1.1.8: Skarlagensfeber |
| | 4 | 8.40 | T4.4.1: Defekt blodplatefunksjon |
| | 5 | 8.10 | T10.3.4: Pneumonier, bakterielle og med ukjent etiologi |
| | 6 | 8.00 | T11.4.4: Halitosis |
| | 7 | 8.00 | T11.3.1.2: Kronisk faryngitt |
| | 8 | 8.00 | T1.7.7: Lymfogranuloma venereum |
| | 9 | 6.90 | T1.11: Bakteriell endokarditt |
| | 10 | 6.40 | T16.9: Kutane bivirkninger av systemiske legemidler |
| 7 | 1 | 23.30 | T8.3.2.2: Hjerteinfarkt med ST-elevasjon |
| | 2 | 22.30 | T21.1.1.1: Nociseptive smerter |
| | 3 | 21.70 | T20.2.1: Akutte smerter |
| | 4 | 17.60 | T20.2.3.2: Bruk av sterkere opioider hos pasienter |
| | 5 | 16.50 | T22.4.1: Postoperativ smertebehandling |
| | 6 | 16.20 | T6.2.3: Spenningshodepine (Tensjonshodepine) |
| | 7 | 15.70 | T22.4.1.3: Opioider i postoperativ smertebehandling |
| | 8 | 15.40 | T20.2.2.1: Praktisk gjennomføring av smertebehandling |
| | 9 | 15.10 | T21.1.1: Smerter |
| | 10 | 14.40 | T20.2: Akutte og kroniske smerter |
| 8 | 1 | 19.30 | T11.3.2.1: Akutt tonsillitt |
| | 2 | 12.90 | T1.5.1: Nedre urinveisinfeksjon |
| | 3 | 12.90 | T1.1.8: Skarlagensfeber |
| | 4 | 12.70 | T1.10: Akutt bakteriell meningitt |
| | 5 | 12.50 | T1.7.5: Syfilis |
| | 6 | 11.30 | T24.2: Nukleærmedisin |
| | 7 | 11.30 | T1: Infeksjonssykdommer |
| | 8 | 10.80 | T1.12: Osteomyelitt |
| | 9 | 10.50 | T1.13: Nekrotiserende fasciitt |
| | 10 | 9.70 | T1.7: Seksuelt overførbare infeksjoner (Soi) |

**Table 4.20:** Task 6 evaluations

| | Precision @ 10 | | | R-precision | | |
|---|---|---|---|---|---|---|
| Case | Task 3 | Task 6 A | **Task 6 B** | Task 3 | Task 6 A | **Task 6 B** |
| 1 | 60% | 50% | 70% | 0.67 | 0.40 | 0.57 |
| 2 | 50% | 40% | 70% | 0.80 | 0.25 | 0.86 |
| 3 | 90% | 40% | 100% | 0.89 | 0.25 | 1.00 |
| 4 | 70% | 70% | 90% | 0.71 | 0.86 | 0.89 |
| 5 | 80% | 50% | 90% | 0.88 | 0.60 | 0.89 |
| 6 | 80% | 30% | 90% | 0.75 | 0.33 | 0.89 |
| 7 | 100% | 50% | 900% | 1.00 | 0.60 | 1.00 |
| 8 | 90% | 30% | 80% | 0.89 | 0.67 | 0.88 |
| Avg | 77.5% | 45.0% | **85.0%** | 0.82 | 0.49 | **0.87** |

**Table 4.21:** Task 6 Kendall tau coefficients

| Case | Task 3 vs Task 6 A | Task 3 vs Task 6 B | Task 6 A vs Task 6 B |
|---|---|---|---|
| 1 | 0.725 | 0.795 | 0.810 |
| 2 | 0.611 | 0.790 | 0.792 |
| 3 | 0.606 | 0.818 | 0.733 |
| 4 | 0.692 | 0.815 | 0.762 |
| 5 | 0.688 | 0.832 | 0.813 |
| 6 | 0.740 | 0.778 | 0.690 |
| 7 | 0.627 | 0.827 | 0.716 |
| 8 | 0.663 | 0.791 | 0.808 |
| Avg | 0.669 | 0.806 | 0.765 |

# CHAPTER 5

# DISCUSSION

In this chapter we discuss task results, limitations to the assignment and look at potential improvements to the system.

## 5.1 Efficient and precise results

The goal of the system was to provide clinicians with an efficient and precise search interface, to help them make therapy recommendations. The correctness of these recommendations is crucial for the system to be trusted by the users.

### 5.1.1 Information retrieval system

To make our information retrieval system fulfill the necessary properties, we had to make some fundamental architectural and design decisions. High performance had to be prioritized to make the system efficient. To achieve this we use the fast and lightweight programming language Python, and we convert all the input files to JSON format. By having files stored as JSON, the loading time is substantially reduced.

As correctness is the most crucial property, we devoted much of our time to tweak algorithms and manually check that the results corresponded to the patient cases. For the two first tasks, we use BM25F, which is a well known algorithm that uses probability to find the most relevant results. In task 3 we make use of a vector model to determine the proximity of the terms to calculate the relevance. In task 6 the probabilistic model and vector model from the prior tasks create a new function to find the most relevant results.

By aggregating a retrieval function based on the algorithms used in task 1,2 and 3, we end up with a system that comply with the initial requirements; efficient and precise. Mathematical ranking techniques like term frequency,

inverse document frequency and term proximity are core parts of the system. These are sound and well known techniques, and used correctly they should give our system a very good therapy recommendation capability.

### 5.1.2 Task results

To see that our results were correct during the work, we did manual evaluations of the results. We are not experts in the medical field, but when reading a case and the suggested result, we could easily determine if the result was relevant or not (at least to a certain degree). After a closer look at the 10 suggested results for each case, we saw that most of the suggested results were relevant.

For task 1 and 2, where the probabilistic method is used, the results depends purely on the summed value of each term included in the query. This approach works, and in most cases the results are relevant. The problem is that terms should not be treated equally, some terms contains more value in the task of classifying than others. For instance the terms "motivert" and "sukkersyke", taken from case 1 sentence 4, should not be thought of as equally significant to the search. We eliminate parts of this problem by removing stopwords from the patient cases before we run the classification. In a perfect world the system would recognize words and understand the meaning/value in a given case, and weight them accordingly to get the perfect results.

For the therapy chapter classification in task 3, we used proximity of the terms to find the most relevant results. This should give more correct classifications than the probabilistic method used in task 1 and 2. Each patient case is compared to all therapy chapters to find the closest match, which is the most relevant result. The importance of individual terms are reduced, and instead all the text in a patient case is used to classify it. Making use of more information should yield better results.

After a manual evaluation of the results from task 3, we can determine the following. Only one case retrieves low amount of relevant therapy chapters, but by looking at the R-precision we can tell that the chapters that actually were relevant must have been among the top results. Probably result 1, 2 and 3. For the other cases, we see the same tendency; R-precision values are high. This suggests that our algorithm for finding the relevant results is good.

It is important to consider how many therapy chapters that actually can be relevant for a case. It might only exist three relevant chapters for case 2, thereby the rest of the results are not relevant. In such cases it is important to have the relevant chapters as top results, which our system seems to handle well.

### 5.1.3 System results

The final system make use of all the methods used to solve the tasks. This is the system we have developed to support clinicians in their work. Evaluating the results from the system give great insight in what level of quality and correctness they can expect from it. If a doctor were to base treatment of a patient on the results of our system, it is crucial to know the reliability of the results. A consequence of incorrect results could lead to wrong treatment and even death.

## 5.2 Limitations

An obvious limitation to the assignment was that only approved libraries could be used to solve the tasks. This point was up for negotiation, and other libraries could be used if approved by the staff. Typically the students would start out by looking at the limited set of choices, and most probably choose one of the suggested. This limitation might have served a good cause; leading the students in to well know libraries that were sure to work for the given tasks. We spend a significant amount of time creating the parser of therapy chapters using only standard Python libraries, while we believe we could have solved this challenge faster using a well-known Python library html5lib[1].

Another limitation was the lack of domain knowledge. When we do manual evaluation of results to a patient case, we can not be completely sure that the results are relevant. It might look relevant to a untrained eye. A domain expert would also discover that relevant results are missing in a result set. Testing in a real environment would overcome this limitation, see section 5.3 for more details.

## 5.3 Potential improvements

This section describes possible improvements to the system. Improvements to fields such as user experience, usability, performance, algorithms and results. This could be accomplished by tweaking of the system on our own or by doing in-field testing with real patients and clinicians. The latter is the most preferred one.

### 5.3.1 Stemmer

We did not use any stemming on the documents, as we were unaware of any known Norwegian stemmer algorithms, but we believe using a stemmer might improve the system.

[1] html5lib: http://code.google.com/p/html5lib/

### 5.3.2 Reference collection

A potential improvement to the system would be to use prior knowledge about patient cases as extra input to our algorithms. Provided with case and correct classification, the system could weight treatments as more probable based on history and similarity. It could well be that a patient case does not give an exact image of what the patient's problem really is, and then it would be beneficial to have a big reference base to help classifying.

### 5.3.3 In-field testing

By doing a field test of the system, we would collect valuable data that could tell us: how the system would be used, how experts want it to be, see how good our results really are, is the system to slow to be used on real cases, is the user interface not intuitive and so on. Experts would get hands-on experience and could quickly determine pros and cons of the system. We believe that this is the most important point in improving the system considerably.

### 5.3.4 More input

If the system was deployed in a real environment, the results from prior cases could be given as extra input to the algorithms. This input could make up a probabilistic model of what that classifications that are often correct, often wrong or where uncertainties often occur. All extra information assist the system in giving the best possible results. Clinicians could record what the system determined as treatment, what they determined as treatment, and then weeks after the treatment see if it was a correct classification or not, and feed this back in to the system.

### 5.3.5 N-grams

Instead of using a bag of words based vector space model, a possible improvement might be to test letter-based n-grams, where each vector cell consist of for example 4 letters found next to each other in documents. Or possibly word-based n-grams like bigrams or trigrams. Such methods would have the benefit of also taking into account proximity of letters and/or words.

# CHAPTER 6

# CONCLUSION

Through several search and retrieval tasks we have learned how algorithms and approaches determine what results we get, and how good these results are. This assignment demonstrates the importance of incorporating multiple mathematic measures to give the best set of relevant results. In addition, we have seen that our information retrieval systems works best if it is provided with several sources of information. We achieved the best result when we combined results from the tasks for the ICD-10 codes, ATC codes and therapy chapters to classify treatment for the patient cases.

A well known management adage is "You can't manage what you don't measure". This implies the importance of evaluating results. As described in the report, we used both manual and automatic evaluation to measure how good the results were. We quickly realized the huge benefits of doing this automatically, but we also understood the difficulties. We believe that automatic evaluation will never be as good as the evaluation done manually by an expert of the domain.

We can conclude that the algorithms we have used are sound, and that they solve the search and retrieval of the tasks in the assignment.

# APPENDIX A

APPENDIX

## A.1 Stopwords

Table A.1 contains a list of Norwegian stopwords used on search queries such as patient cases and therapy chapters. We found started with an initial list of stopwords that we found online[1]. We added additional words with low relevance that are frequently used in patient cases.

## A.2 Medical terms

Table A.2 list medical terms which we found in patient cases. These are used for automatically evaluating if a search result is relevant or not.

## A.3 Patient cases

This chapter contains patient cases used as input in this project. Norwegian stop words have been removed from these patient cases.

[1] Stopword source: http://www.wisweb.no/999/147/33899-170.html

**Table A.1:** Norwegian stopwords

| A - D | D - H | H - K | K - N | N - S | S - Å |
|---|---|---|---|---|---|
| alle | dit | har | kun | noe | so |
| andre | ditt | hennar | kunne | noen | som |
| at | du | henne | kva | noka | somme |
| av | dykk | hennes | kvar | noko | somt |
| bare | dykkar | her | kvarhelst | nokon | start |
| begge | då | hit | kven | nokor | stille |
| behandling | eg | hjå | kvi | nokre | syk |
| ble | ein | ho | kvifor | ny | så |
| blei | eit | hoe | lage | nå | sånn |
| bli | eitt | honom | lang | når | tid |
| blir | eller | hos | lege | og | til |
| blitt | elles | hoss | lik | også | tilbake |
| bort | en | hossen | like | om | um |
| bra | ene | hun | man | opp | under |
| bruk | eneste | hva | mange | oss | upp |
| bruke | enhver | hvem | me | over | ut |
| både | enn | hver | med | pasienten | uten |
| båe | er | hvilke | medan | pasienter | var |
| bør | et | hvilken | meg | pga | vart |
| ca | ett | hvis | meget | på | varte |
| da | etter | hvor | mellom | rett | ved |
| de | for | hvordan | men | riktig | verdi |
| deg | fordi | hvorfor | mens | samme | vere |
| dei | forsøke | i | mer | seg | verte |
| deim | fra | ikke | mest | selv | vi |
| deira | fram | ikkje | mg | si | vil |
| deires | få | ingen | mi | sia | ville |
| dem | før | ingi | min | sidan | vite |
| den | først | inkje | mine | siden | vore |
| denne | gjorde | inn | mitt | sin | vors |
| der | gjøre | innen | mot | sine | vort |
| dere | god | inni | mye | sist | vår |
| deres | gå | ja | mykje | sitt | være |
| det | går | jeg | må | sjøl | vært |
| dette | ha | kan | måte | skal | å |
| di | hadde | kom | ned | skulle | |
| din | han | korleis | nei | slik | |
| disse | hans | korso | no | slutt | |

**Table A.2:** Medical terms

| A - H | H - P | P - V |
|---|---|---|
| acetonlukt | hemoglobin | penicillin |
| allergier | hemolytiske | penicillintabletter |
| analgetika | hemorroider | peroral |
| angina | hjernehinnebetennelse | petekkier |
| antibiotika | hjertelydene | pharynx |
| apocillin | hodepine | postoperativt |
| astma | hostet | pulmicort |
| atrovent | influensa | pulsen |
| attføring | injiserer | pusteplager |
| avføring | insulin | rektaleksplorasjon |
| avføringen | insulinkrevende | respirator |
| bakteriologisk | insulinpenn | røntgenbildet |
| blod | intramuskulære | sacrum |
| blodkulturer | intravenøs | sederende |
| blodmangel | intravenøse | serogruppe |
| blodprøver | intravenøst | serumspeil |
| blodprøves | intuberes | skjelett |
| blodtrykk | ketorax | slim |
| blodtrykket | kloramfenikol | smertelindring |
| blødningskilde | kneplager | småblødninger |
| bricanyl | kortisonpreparat | spinalpunksjon |
| bronkiene | krampeanfall | spinalvæske |
| brystsmerter | kvalm | spinalvæsken |
| cancer | lungeflatene | spirometri |
| colli | lymfeknuter | spirometriundersøkelse |
| colonoscopi | mellitus | stetoskop |
| desorientert | meningitidis | strept.test |
| diabetes | meniskoperert | streptokokker |
| diazepam | meniskplager | streptokokktonsilitt |
| diplokokker | metastaser | støtdoser |
| dolcontin | mikrobiologisk | submandibulært |
| ekspiratoriske | morfin | sukkersyke |
| ekspiriet | muskulatur | svelgbesvær |
| endetarmsåpningen | muskulært | tarmveggen |
| endoskopisk | nakkestiv | tonsilleforstørrelse |
| femoris | napren | tonsillene |
| foetor | neisseria | tonsiller |
| fractura | nevrogen | totalprotese |
| gane | nitroglycerin | tykktarmen |
| glandler | opioider | vaksinasjon |
| halebenet | ortopedisk | væsketillførsel |
| halsflora | palpasjonsømhet | |
| halsinfeksjon | paralgin | |
| halsprøve | pectoris | |

**Table A.3:** Patient case 1

| # | Lines (stopwords removed) |
|---|---|
| 1 | Eva Andersen skoleelev hatt insulinkrevende diabetes mellitus 3 år |
| 2 | bror diabetes brukt insulin flere år |
| 3 | bruker insulinpenn hurtigvirkende insulin injiserer huden hvert måltid dessuten injeksjon langtidsvirkende insulin kvelden |
| 4 | lite motivert håndtere sukkersyke |
| 5 | uteblir kontroller, delvis tatt insulin periodevis ignorert kostråd |
| 6 | synes leit fått sukkersyke |
| 7 | Eva siste døgn følt tungpusten, kvalm kastet opp |
| 8 | Rødmusset |
| 9 | Hurtig pust |
| 10 | Tørr pannen |
| 11 | Acetonlukt |
| 12 | Normalt blodtrykk |
| 13 | delvis uklar, vurderer henvisning sykehus |

**Table A.4:** Patient case 2

| # | Lines (stopwords removed) |
|---|---|
| 1 | 56 år gammel mann stort sett frisk tidligere bortsett meniskplager ført meniskoperert knær |
| 2 | arbeidet avløser jordbruk/skogbruk |
| 3 | attføring kneplager |
| 4 | Lungepoliklinikken hatt tungpust siste 3 - 4 månedene |
| 5 | tungpust hvile, tung pusten bakker trapper , stoppe hvile gått 400 – 500 meter flat mark, kortere distanser |
| 6 | tillegg tungpust hatt følelsen tett brystet |
| 7 | hostet mye, fått del slim grønt farge føler "der noe" brystet |
| 8 | lagt merke "skrål piping" brystet |
| 9 | symptomene hatt omtrent lenge følt verre pusten |
| 10 | nærmere utspørring kommer nok tyngre pusten siste par år |
| 11 | merket fått problemer holde følge jevnaldringer situasjoner tidligere gått greitt festet spesielt brakt bane |
| 12 | Imidlertid inntrådt merkbar forverring pusten siste 3 – 4 måneder |
| 13 | aldri plaget astma allergier tidligere livet |
| 14 | røkt ca. 40 år, mesteparten tiden 3 pakker rulletobakk uka, reduserte pakke uka par år sluttet helt 3 måneder pusteplager |
| 15 | undersøkelsen Lungeavdelingen ubesværet pusten hvile |
| 16 | lytting stetoskop lungeflatene høres unormale lyder ("pipelyder") puster (i ekspiriet), mindre grad puster inn |
| 17 | Hjertelydene normale |
| 18 | Pulsen 80 regelmessig |
| 19 | Røntgenbildet lungene normalt |
| 20 | gjort spirometri viser VK (vitalkapasiteten volumet luft maksimalt fylle lungene pust) 2,56 liter (53% normalt) FEV1 (det forserte ekspiratoriske volum volumet blåse lungene sekund) 1,28 liter (34% normalt) |
| 21 | gitt antibiotika behandlingen bricanyl Atrovent (som utvider bronkiene ) pulmicort (et kortisonpreparat inhalerer) hvert betydelig bedre idet pusten lettere, hosten bedre grønnfargete oppspyttet forsvant gjenvant vante tilstand |
| 22 | spirometriundersøkelse viste måned VK 3,20 liter (66% normalt) FEV1 2,02 liter (53% normalt), altså betydelig bedring |

**Table A.5:** Patient case 3

| # | Lines (stopwords removed) |
|---|---|
| 1 | Hanne 9. klasse ungdomsskolen, to yngre søsken |
| 2 | tidligere betydning, aldri hospitalisert |
| 3 | morgen februar klager Hanne føler syk, litt vondt halsen, hodepine smerter hele kroppen temperatur 38,7 °C |
| 4 | del influensa distriktet tror foreldrene Hanne holder utvikle |
| 5 | holder hjemme skolen dårligere utover dagen |
| 6 | far kommer hjem 16-tiden, Hanne litt omtåket, ligger døs temperaturen måles 40,9 °C |
| 7 | undersøkelsen finner legen desorientert, nakkestiv spredt kroppen petekkier ("småblødninger" huden) |
| 8 | Blodtrykket måles 105/70 mm Hg |
| 9 | Hanne innlegges umiddelbart Barneklinikken mistanke "smittsom hjernehinnebetennelse" |
| 10 | mottagelsen sykehuset Hanne undersøkt vakthavende lege, henges intravenøs væsketillførsel, blodprøves takes gjøres spinalpunksjon |
| 11 | Spinalvæsken tydelig blakket |
| 12 | 2 blodkulturer tatt, startes umiddelbart intravenøs antibiotika Penicillin G Kloramfenikol |
| 13 | par timer innkomsten får Hanne par minutters krampeanfall, behandles diazepam intravenøst |
| 14 | besluttes Hanne intuberes legges respirator |
| 15 | Dagen innkomst rapporterer mikrobiologisk laboratorium funn Gram-negative diplokokker spinalvæske blodkulturer, kommer oppvekst Neisseria meningitidis serogruppe C |
| 16 | Hannes to yngre søsken får penicillintabletter uke (Helsedirektoratets forskrift) |
| 17 | Kommunelegen setter igang vaksinasjon hele Hannes familie, nære venner skoleklasse |

**Table A.6:** Patient case 4

| # | Lines (stopwords removed) |
|---|---|
| 1 | Trond Øvrebotten, 42 år, oppsøker fastlegen 2-3 måneder følt ubehag, trykk brystet anstrengelse |
| 2 | gang kjent opphisselse |
| 3 | røyker 10-15 sigaretter daglig, trener spiser sunn mat |
| 4 | far døde plutselig ca. 50 år gammel |
| 5 | Legen starter medikamentell pasientens symptomer angina pectoris henviser spesialistundersøkelse |
| 6 | rekker komme innlegges sykehuset grunn kraftige brystsmerter litt bedre nitroglycerin, helt bort |

**Table A.7:** Patient case 5

| # | Lines (stopwords removed) |
|---|---|
| 1 | 64 år gammel kvinne tidligere stort sett frisk |
| 2 | siste 5 månedene merket følelse ufullstendig tømming avføring |
| 3 | flere anledninger spor friskt blod avføringen |
| 4 | tilskriver hemorroider hatt før, ventet over |
| 5 | egen finner galt vanlig undersøkelse |
| 6 | rektaleksplorasjon (kjenne endetarmsåpningen finger) kjenner legen vidt kanten uregelmessighet tarmveggen |
| 7 | Legen ser spor gamle ytre hemorroider sannsynlig blødningskilde nå |
| 8 | Prøver blod avføringen positive |
| 9 | Blodprøver viser lett blodmangel (hemoglobin 10.8; normalt kjønn alder >12) |
| 10 | Øvrige blodprøver normale |
| 11 | henvist colonoscopi (endoskopisk undersøkelse tykktarmen) |

**Table A.8:** Patient case 6

| # | Lines (stopwords removed) |
|---|---|
| 1 | første konsultasjon funnet forstørrede tonsiller hvitlig belegg forstørrede glandler sider halsen |
| 2 | tatt halsprøve strept.test positiv |
| 3 | fremkom kjæresten nylig gjennomgått streptokokktonsilitt |
| 4 | startet peroral penicillin (Apocillin) vanlig dosering |
| 5 | kommer konsultasjon manglende effekt behandlingen penicillin |
| 6 | verre, fikk svelgbesvær, fikk fast føde, besværligheter væske |
| 7 | Samtidig vedvarende slapp dårlig allmenntilstand |

**Table A.9:** Patient case 7

| # | Lines (stopwords removed) |
|---|---|
| 1 | Berta fikk påvist cancer mamma våren fem år siden |
| 2 | påvist metastaser skjelett innlagt ortopedisk avdeling fractura colli femoris behandlet innsetting totalprotese |
| 3 | hensyn postoperativt forløp opptrening hoften beskrives minste problem |
| 4 | hensyn smerter sier uttalt halebenet |
| 5 | smerter høyere opp, tilsvarende os sacrum |
| 6 | stemmer godt ferske funn rtg. bekken |
| 7 | Smertene føles trykkende karakter |
| 8 | utstrålende smerter tegn overfølsomhet overliggende hudområde |
| 9 | spesiell palpasjonsømhet muskulatur |
| 10 | Således holdepunkter nevrogen muskulært betingede smerter |
| 11 | svært tilbakeholden forbruk analgetika |
| 12 | Bruker Napren E 250 x 2 samt Paracet behov |
| 13 | enig fortsetter Napren E uendret dose |
| 14 | angir intoleranse Paralgin Forte, unngår moderat virkende opioider gir heller foreløpig Ketorax 5 behov |
| 15 | Forklarer lav terskel ta Ketorax |
| 16 | justere enkeltdosene 1/2-2 tabletter, føler gir effekt uakseptable bivirkninger |
| 17 | Legger vekt smertelindring gi vesentlig økt livskvalitet |
| 18 | tidligere prøvd morfin, svært kvalm dette |
| 19 | Sannsynligvis tolerere «fast serumspeil» morfin form Dolcontin; tross kvalm intramuskulære intravenøse støtdoser |
| 20 | orientert forholdsvis kort utvikler toleranse sederende virkning opioider |

**Table A.10:** Patient case 8

| # | Lines (stopwords removed) |
|---|---|
| 1 | Thomas 8. klasse aktiv gutt fotball volleyball |
| 2 | par dager vondt halsen, dårlig allmenntilstand feber tilkaller foreldrene legevakt |
| 3 | Foreldrene frykter Thomas halsinfeksjon lurer antibiotika (Thomas reise leirskole forbindelse konfirmasjons-forberedelsene 3 dager) |
| 4 | undersøkelse pharynx finner legen tonsillene forstørrete, ses hvitlig belegg tonsillene |
| 5 | hovne lymfeknuter submandibulært |
| 6 | Legen besluttet starte penicillin-behandling (tabletter Apocillin 0,5 + 0,5 + 1 mill IE) tatt halsprøve innsendes bakteriologisk dyrking |
| 7 | Tredje dag legevaktbesøk Thomas fortsatt omkring 39°C kommer undersøkelse legekontoret |
| 8 | undersøkelse finnes økende tonsilleforstørrelse - møtes nesten midtlinjen |
| 9 | Tonsillene belagt, ses flere petekkier (småblødninger) bløte gane |
| 10 | stygg foetor ex ore |
| 11 | puster gjennom munnen tett nesen |
| 12 | Telefonhenvendelse mikrobiologisk laboratorium avklarer innsendte halsprøve viser moderat hemolytiske streptokokker gr C blandet halsflora |