



Generierung von Diagnosehypothesen für die DRG-Kodierung durch semantische Analyse der Medikation in Arztbriefen

Thorsten Seddig^{1,2}, Stefan Schulz¹

¹Universitätsklinikum Freiburg, Medizinische Informatik

²Averbis GmbH, Freiburg



- Problemstellungen und Zielsetzungen des Pilot-Projektes
- Elemente einer semantischen Analyse von Arztbriefen für die DRG-Klassifizierung
- Lokalisierung von Textpassagen mit Medikamentennamen
- Verknüpfung mit DRG-Groupen
- Präsentation der Kandidaten-ICDs
- Evaluierung



- Leistungsabbildung der stationären Versorgung über das **DRG-System** (Diagnosis Related Groups)
- DRG → **Klassifikation von Behandlungsfällen**
 - aufgrund von Hauptdiagnose und Nebendiagnosen (ICD)
 - Prozeduren (OPS)
 - Alter
 - Geschlecht
 - Komplikationen (Liegendauer)
- Aufgabe des **Medizinisches Controllings** ist die Sicherstellung der Korrektheit und Vollständigkeit des Klassifikationskodes.
 - Erfordert aufwändige manuelle Durchsicht der Akten



- Unvollständige Diagnosedokumentation führt zu **Unterkodierung**
 - **geringere Leistungsvergütung** aufgrund geringerer DRG-Einstufung
 - **Hypothesen:**
Vorhandene Medikation → Diagnose(n)
Auf das wahrscheinlichste Medikations/Diagnose-Paar kann durch maschinelles Lernen geschlossen werden
 - **Ziel:**
Identifizierung von **fehlenden** durch **Abgleich** mit vorhandenen Diagnosen
 - **Ziel:**
Minimierung der Rate an Unterkodierung (Mindereinnahmen) und erleichternde manuelle Arbeit im Medizincontrolling



- **Verordnete Medikamente** in Arztbriefen
- **Kodierte Diagnosen** aus dem Klinikinformationssystem
- **Beziehungen** zwischen den Medikamentenverordnungen und Diagnosen
 - Identifizierung von freitextlichen **Medikamentenverordnungsmuster**
 - **Abbildung der Beziehungen** durch wahrscheinlichkeitstheoretische Modelle (**Bayes-Theorem**)
 - **Instanziierung der Modelle** durch Ermittlung der Häufigkeiten der aufgetretenen Medikamente und Diagnosen.
 - **Zielsetzung:**
Ableitung von Diagnosen aufgrund vorhandener Medikation der Patienten → Identifizierung von fehlenden Diagnosen



- Untermenge von Behandlungsfällen mit **bekanntermaßen guter Kodierung** von nichtchirurgischen Diagnosen:
 - Innere Medizin
 - Neurologie
- **Trainingsdatensätze:**
 - **Metadaten:** ICD, OPS, Alter, Geschlecht aus **KIS**
 - Medikamentendaten (**Wirkstoffe**), extrahiert aus Arztbriefen
- **Reduktion des ICD-Merkmalraums:**
 - Ignorieren von Codes die keine DRG-Relevanz besitzen



- **Instanziierung / Training**
 - Anzahl an annotierten Arztbriefen (Diagnosen/Medikamente)
 - $P(\text{ICD1} | \text{Med1})$, $P(\text{ICD1} | \text{Med1}, \text{Med2})$? $P(\text{Med1})$, $P(\text{Med2})$
 - Vorteil gegenüber Fachinformationen
→ Wahrscheinlichkeiten + Kombinationen

- **Beispiel**

Wirkstoff	Diagnose	Häufigkeit
ASS	Kopfschmerz	6000
Benazepril	Hypertonie	500
ASS, Benazepril	KHK	10
ASS, Benazepril	Kopfschmerz	5

Demographische Daten

Aufnahmedatum, Entlassdatum, Geburtsdatum, Geschlecht, Verweildauer ...

Diagnosen (ICD-10 Version GM 2007)

- 1 K43.9 Hernia ventralis ohne Einklemmung und ohne Gangrän
- 2 C21.1 Bösartige Neubildung: Analkanal
- 3 I10.90 Essentielle Hypertonie, nicht näher bezeichnet: Ohne Angabe einer hypertensiven Krise
- 4 Z88.0 Allergie gegenüber Penicillin in der Eigenanamnese

Prozeduren (OPS-301 Version GM 2007)

- 1 5-536.44 21.03.2007 Verschluss einer Narbenhernie: Mit alloplastischem Material: Offen chirurgisch als Bauchwandersatz
- 2 8-915 21.03.2007 Injektion eines Medikamentes an andere periphere Nerven zur Schmerztherapie

Grouping-Ergebnisse (Fallpauschalenkatalog G-DRG Version 2007)

DRG: G08B Komplexe Rekonstruktion der Bauchwand, Alter > 0 Jahre, ohne äußerst schwere CC

MDC: 06 Krankheiten und Störungen der Verdauungsorgane

PCCL: 3 schwere CC

Status: 00 Normale Gruppierung

Basisrate: 2.907,96 €

Basisbetrag:	3.390,68 €	Rel.-Gew.:	1,166	1. Tag Abschlag:	1
Zu-/Abschlag:	0,00 €	Eff.-Gew.:	1,166	1. Tag Zuschlag:	14
Entgelt (inkl. Zu-/Ab.):	3.390,68 €	Pat.-Status:	Normallieger	Mittlere Verw.:	7,5



Sehr geehrter Herr Kollege,

...

Diagnosen:

- Narbenhernie
- Z. n. transanaler Vollwandexzision des Malignen Melanoms im Analkanal 10/2006
- Z. n. villösem Adenom Sigma ED 01/99, sequentielle endoskopische Abtragung 01/1999 bis 06/2000
- Z. n. wiederholten Abtragungen Sigmapolypen
- Z. n. Polypenknospenabtragung Coecum 11/2001
- Mamma-Karzinom rechts ED 1969
- Z. n. Ablatio Mammae re.
- Z. n. Radiotherapie und Adnexektomie 1969
- chronisches Lymphödem rechter Arm
- rez. Erysipel (1990, 1994, 1997, 1998, 2000, 2002)

Medikation bei Entlassung:

- **Madopar** 62,5mg 1-0-1
- Movicol Btl. bei Bedarf
- Pantozol 40 mg 0-0-1
- Clexane 0,2 ml sc. 1 x die (bis zur Entlassung)
- Noctamid 1 mg bei Bedarf
- Novalgin 40 gtt. bei Bedarf

Demographische Daten

Aufnahmedatum, Entlassdatum, Geburtsdatum, Geschlecht, Verweildauer ...

Diagnosen (ICD-10 Version GM 2007)

- 1 K43.9 Hernia ventralis ohne Einklemmung und ohne Gangrän
- 2 C21.1 Bösartige Neubildung: Analkanal
- 3 I10.90 Essentielle Hypertonie, nicht näher bezeichnet: Ohne Angabe einer hypertensiven Krise
- 4 Z88.0 Allergie gegenüber Penicillin in der Eigenanamnese
- 5 G20.90 Primäres Parkinson-Syndrom, nicht näher bezeichnet: Ohne Wirkungsflektuation

Prozeduren (OPS-301 Version GM 2007)

- 1 5-536.44 21.03.2007 Verschluss einer Narbenhernie: Mit alloplastischem Material: Offen chirurgisch als Bauchwandersatz
- 2 8-915 21.03.2007 Injektion eines Medikamentes an andere periphere Nerven zur Schmerztherapie

Grouping-Ergebnisse (Fallpauschalenkatalog G-DRG Version 2007)

DRG: G08A Komplexe Rekonstruktion der Bauchwand, Alter > 0 Jahre, mit äußerst schweren CC

MDC: 06 Krankheiten und Störungen der Verdauungsorgane

PCCL: 4 äußerst schwere CC

Status: 00 Normale Gruppierung

Basisrate: 2.907,96 €

Basisbetrag:	5.667,61 €	Rel.-Gew.:	1,949	1. Tag Abschlag:	3
Zu-/Abschlag:	0,00 €	Eff.-Gew.:	1,949	1. Tag Zuschlag:	26
Entgelt (inkl. Zu-/Ab.):	5.667,61 €	Pat.-Status:	Normallieger	Mittlere Verw.:	13,0



1. Fachinformationen

→ Liste potentieller Diagnosen

2. Arztbrief-Statistik

→ Nach Wahrscheinlichkeiten sortierte Liste potentieller Diagnosen



- **Extraktion von Medikamentennamen**

▪ **Mapping auf strukturiertes Vokabular: ...**

- Standardterminologie
- Morphosaurus-Äquivalenzklassen
- UMLS

▪ **Substanznamen** und entsprechende **Handelsnamen** werden auf dieselben **Id's** abgebildet

- Identifikation von **Medikamentennamen in Freitextpassagen** durch
- **Pattern (String) Match** (Heuristiken, OpenNLP-Tools)
 - **MorphoSaurus-Deskriptoren**

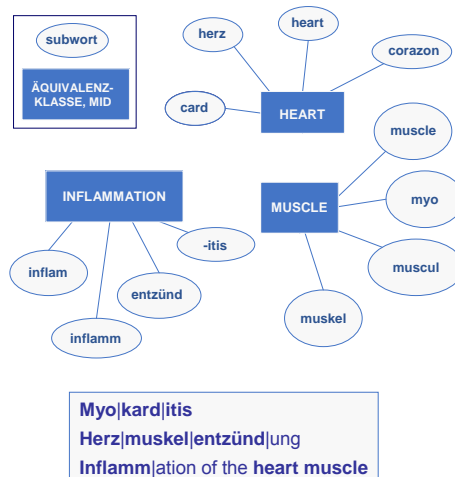


Sprachliche Variationen der natürlichen Sprache:

- **Flexion:** Erkrankungen, Bluthochdrucks, Risiken
- **Derivation:** Krankheit, diastolisch, leukozytär
- **Komposition:** Gemeinde|krank|en|pflege|station|en
- **Akronyme:** AIDS, SARS, OECD
- **Wortstellung:** Nierenzyste, Zystenniere
- **Eigennamen:** Aspirin, ASS, ...



- Konventionelle Fachlexika enthalten rund 200.000 Einträge
- Wortschatz um das 7-fache größer !
- ✓ Innovation: **Subwortlexika**
- ✓ sprachübergreifend verlinkt über **Thesaurus**
- Datenreduktion um 90%
- Berücksichtigung aller sprachlichen Varianten





High TSH values suggest the diagnosis of primary hypothyroidism ...

Erhöhte TSH-Werte erlauben die Diagnose einer primären Hypothyreose ...

Original



High TSH values suggest the diagnosis of primary hypothyroidism ...

Erhöhte TSH-Werte erlauben die Diagnose einer primären Hypothyreose ...

Original

Orthographische
Normalisierung

Orthographische
Regeln

high tsh values suggest the diagnosis of primary hypothyroidism ...

erhoehte tsh-werte erlauben die diagnose einer primaeren hypothyreose ...





High TSH values suggest the diagnosis of primary hypothyroidism ...

Orthographische Normalisierung

high tsh values suggest the diagnosis of primary hypothyroidism ...

Erhöhte TSH-Werte erlauben die Diagnose einer primären Hypothyreose ...

Orthographische Regeln

erhoehte tsh-werte erlauben die diagnose einer primaeren hypothyreose ...

Original



Morph. Segmentierung
Subwort Lexikon

high tsh value s suggest the diagnosis of primar y hypo thyroidism

er hoeh te tsh wert e erlaub en die diagnos e einer primaer en hypothyre ose



High TSH values suggest the diagnosis of primary hypothyroidism ...

Orthographische Normalisierung

high tsh values suggest the diagnosis of primary hypothyroidism ...

Erhöhte TSH-Werte erlauben die Diagnose einer primären Hypothyreose ...

Orthographische Regeln

erhoehte tsh-werte erlauben die diagnose einer primaeren hypothyreose ...

Original

MID-Repräsentation

UP TSH VALUE SUGGEST
DIAGNOST PRIMAR SMALL
THYREOS

Semantische Normalisierung

high tsh value s suggest the diagnosis of primar y hypo thyroidism

UP TSH VALUE PERMIT
DIAGNOST PRIMAR SMALL
THYREOS

Subwort Thesaurus

er hoeh te tsh wert e erlaub en die diagnos e einer primaer en hypothyreos



Morph. Segmentierung
Subwort Lexikon



Typ	Suche nach... Mapping von...	Findet... nach...	Sonstige Methoden	Morpho- Saurus
Tippfehler:	Brutskrebs	Brustkrebs	✓	✓
Rechtschreibfehler:	Appoplex	Apoplex	✓	✓
Phonetik:	Syfilis	Syphilis	✓	✓
Singular/Plural:	Risiko	Risiken	?	✓
Wortarten:	leukozytär	Leukozyt	?	✓
Vertauschungen:	Chronische Bronchitis	Bronchitis, chronisch	?	✓
Zusammenhänge:	Brustkrebs	Karzinom des Brustdrüsengewebes	---	✓
Wortbildung:	Darmkrebsrisikoreduzierung	Reduktion des Risikos von Darmkrebs	---	✓
Synonyme:	Schlaganfall	Apoplex	---	✓
Abkürzungen:	WHO	Weltgesundheitsorganisation	---	✓
Laie/Experte:	Darmkrebs	Kolonkarzinom	---	✓
Mehrsprachigkeit:	Herzmuskelentzündung	Inflammation of the heart muscle, myocarditis	---	✓

→ Steigerung der Retrieval-Performanz auf 150% (einsprachig) und 120% (mehrsprachig, MedInfo 2007)



- **G-DRG Grouper Software**
 - Ermittlung der DRG-Relevanz von potentiell fehlenden Diagnosen
 - Simulation der neuen DRG-Einstufung



- Aufbau einer **Web-Interface Unterstützung**
- Eingegebene Textpassagen aus dem Arztbrief
 - **Ranglisten der Diagnosen** orientiert an der Wahrscheinlichkeit für eine Diagnose
- Der **verantwortliche Text** für die verordneten Medikamente wird mit einem **Marking** annotiert
- Dient dem Medizincontroller als **Entscheidungsunterstützung**



- **Zufallsstichproben-Evaluierung**
100 Fälle aus verschiedenen klinischen Fachabteilungen werden von zwei medizinischen Controllern begutachtet
 - Jeder Controller überprüft **Cross-Over** 50 Fälle mit und ohne die Web-Basierte Unterstützung
- **Gemessen werden:**
 - Die für jede Fallprüfung benötigte Zeit
 - Die Anzahl der zusätzlich gefundenen Diagnosen
 - Der zusätzlich abzurechnende Erlös nach dem DRG-System
- **Korrektheit und Vollständigkeit**
- **DRG-Relevanz:**
 - Rate der höherwertigen DRG-Einstufungen



- Zeitplan
- Gerade begonnen . . .
- Proof-of-Concept (inkl. Evaluation) bis Anfang 2009