

Merkmalsextraktion aus klinischen Routinedaten mittels Text Mining

Preprint – to appear in OPHTHALMOLOGE

[Feature extraction from clinical routine data using text mining]

Bastian Grundel¹, Marc-Antoine Bernardeau¹, Holger Langner², Christoph Schmidt³,
Daniel Böhringer⁵, Marc Ritter², Paul Rosenthal³, Andrea Grandjean⁴, Stefan Schulz⁴,

Philipp Daumke^{4**}, Andreas Stahl^{1*}

¹Klinik und Poliklinik für Augenheilkunde, Universitätsmedizin Greifswald;

²Professur Medieninformatik, Hochschule Mittweida;

³Institute for Visual and Analytic Computing, Universität Rostock;

⁴Averbis GmbH, Freiburg;

⁵Klinik für Augenheilkunde, Universitätsklinikum Freiburg,
Medizinische Fakultät, Universität Freiburg

#Korrespondierender Autor: Dr. med. Philipp Daumke, Salzstraße 15, 79098 Freiburg
0049 761 7083940, philipp.daumke@averbis.com

*Diese Autoren haben zu gleichen Teilen beigetragen

Zusammenfassung

Hintergrund: Anti-VEGF-Medikamente prägen heute die Therapie von Makulaerkrankungen. In diesem Zusammenhang wird eine Fülle zusätzlicher Daten erhoben. Damit ließen sich Behandlungsverläufe besser verstehen und vorhersagen. Allerdings sind diese Informationen meist nur in freitextlicher Form verfügbar.

Ziel der Arbeit: Wie weit auswertbare Information aus Kliniktexten automatisch gewonnen werden kann, sollte in einer retrospektiven Studie analysiert werden. Ziel war die Einschätzung der Eignung eines zu diesem Zweck parametrisierten Text-Mining-Verfahrens.

Material und Methoden: Es standen Daten zu 3.683 Patienten zur Verfügung, davon 40.485 Arztbriefe. Für einen Teil waren die interessierenden Daten (Visus, Tensio und Begleitdiagnosen) auch strukturiert erfasst worden und konnten so als Goldstandard für die Textanalyse dienen. Diese wurde mit dem System Averbis Health Discovery durchgeführt. Zur Optimierung auf die Extraktionsaufgabe wurde dieses mit Regelwissen sowie mit einem deutschsprachigen Fachvokabular für die internationale Medizinterminologie SNOMED CT angereichert.

Ergebnisse: Die Übereinstimmung der Datenextrakte mit den strukturierten Datenbankeinträgen wird durch den F1-Wert beschrieben. Hierbei ergab sich eine Übereinstimmung von 94,7 % für den Visus, 98,3 % für die Tensio, und 94,7 % für begleitende Diagnosen. Die manuelle Analyse nicht übereinstimmender Fälle zeigte zur Hälfte, dass Textinhalte aus verschiedenen Gründen von Datenbankinhalten abwichen. Nach einer daraus berechneten Adjustierung lagen die F1-Werte noch 1 bis 3 % über den zuvor ermittelten Werten.

Diskussion: Für den betrachteten Arztbriefkorpus und die beschriebene Fragestellung sind Text-Mining-Verfahren sehr gut geeignet, um Inhalte zur weiteren Auswertung strukturiert aus Kliniktexten zu extrahieren.

Schlüsselwörter

- Makuladegeneration
- Verarbeitung natürlicher Sprache
- SNOMED CT
- Elektronische Patientenakten
- Klinische Entscheidungsunterstützung

Abstract

Background: Anti-VEGF drugs are currently used to treat macular diseases. This has led to a wealth of additional data, which could help understand and predict treatment courses. However, this information is usually only available in free text form.

Objectives: A retrospective study was designed to analyse how far interpretable information can be obtained from clinical texts by automated extraction. The aim was to assess the suitability of a text mining method that was customised for this purpose.

Material and Methods: Data on 3,683 patients were available, including 40,485 discharge letters. Some of the data of interest (Visual acuity (VA), intraocular pressure (IOP) and accompanying diagnoses) were not only recorded textually but also entered in a database and could thus serve as a gold standard for text analysis.

Test was analysed using the Averbis Health Discovery text mining platform. To optimize for the extraction task, rule knowledge and a German value set linked to the international terminology standard SNOMED CT was manually added.

Results: The correspondence between extracted data and the database entries is described by the F1 value. There was an agreement of 94.7% for VA, 98.3% for IOP, and 94.7% for the accompanying diagnoses. Manual analysis of inconsistent cases showed in 50% that text content did not match the database content for various reasons. After an adjustment, F1 values 1 - 3% above the previously determined values were obtained.

Conclusion: Text mining methods are very well suited for the considered clinical text corpus and the problem described, in order to extract content from clinical texts in a structured manner for further evaluation.

Keywords (MeSH)

- Macular Degeneration
- Natural Language Processing
- Systematized Nomenclature of Medicine
- Electronic Health Records
- Decision Support Systems, Clinical

Kurze Hinführung zum Thema

Mit der elektronischen Verfügbarkeit von klinischen Routinedaten besteht wachsendes Interesse an Sekundärnutzungsszenarien [01][02][03], in erster Linie für wissenschaftliche Auswertungen. Der unmittelbare Zugang zu diesen Daten ist jedoch dadurch erschwert, dass sie in klinischen Texten, wie Befundberichten, OP-Berichten oder Arztbriefen verborgen sind. Verfahren des Text Mining, also der inhaltlichen Erschließung von Textinhalten mittels spezialisierter Software, kommt daher eine wichtige Bedeutung zu.

Hintergrund

Die Therapie der altersabhängigen Makuladegeneration (AMD), einer häufigen Ursache für bleibende Sehminderung bis hin zur Erblindung, wurde in den letzten 15 Jahren durch die Injektion von Inhibitoren des VEGF (Vascular Endothelial Growth Factor) in den Glaskörper revolutioniert. Ähnliches gilt für die Behandlung des diabetischen Makulaödems (DMÖ) und des Makulaödems nach retinalem Venenverschluss (RVO). Ein häufiges Problem ist die exakte Identifizierung dieser Indikationen durch die Analyse von Krankenakten, z.B. um Prävalenzen und Inzidenzen in *Real-Life*-Kohorten zu berechnen oder Machbarkeitsanfragen (*Feasibility Questionnaires*) für klinische Studien zu beantworten.

Durch die elektronische Dokumentation hat sich vielerorts ein "Datenschatz" angesammelt, der zahlreiche Jahrgänge von tausenden Patienten umfasst. Eine unmittelbare Verwendung dieser Daten ist jedoch selten möglich, da sie sich in gering strukturierten klinischen Texten, wie z.B. Arztbriefen verbergen. Fortschritte bei der Analyse natürlicher Sprache, hier insbesondere im Bereich des Text Mining [04][05], ermöglichen es zunehmend, Daten guter Qualität für die weitere Analyse aus klinischen Texten zu gewinnen [06].

Das BMBF-Projekt TOPOs¹ hat zu 3.683 Patienten mit AMD, DMÖ und RVO 40.485 unstrukturierte und anonymisierte Datensätze aus den Jahren 2007 bis 2017 analysiert. Ziel war, Diagnosen und diagnostische Merkmale mittels Text Mining zu extrahieren, um Prävalenz und Inzidenz zu bestimmen, sowie Vorhersagen zu Therapieverlauf und -Response zu treffen. Die Text-Mining-Resultate speisen einen Kohortenfinder, der für Planung von Therapiestudien notwendige Abschätzungen erlaubt. Beispielhaft werden hier die Ergebnisse zur Kohortenabschätzung von Patienten mit unbehandelter (*treatment naive*) chorioidaler

¹ <http://topos.averbis.de/>

Neovaskularisation (CNV) bei AMD vorgestellt. Das Verfahren wurde im Kontext des TOPOS-Projekts evaluiert, ist aber prinzipiell für andere ophthalmologische Studien anwendbar.

Material und Methoden

Datengrundlage

Die Daten stammen aus der Klinik für Augenheilkunde des Universitätsklinikums Freiburg. Eingeschlossen wurden Patienten, bei denen ab 2013 bis zum Erhebungszeitraum Anfang 2018 eine anti-VEGF Therapie indiziert und begonnen, bzw. bei früherer Indikation noch durchgeführt wurde. Die Klinik für Augenheilkunde Freiburg behandelt Patienten in der Regel nach einem (adaptierten) PRN-Schema Serien von drei Injektionen.

Die Datensätze bestanden neben unstrukturierten Texten aus strukturierten Inhalten (Tabelle 1). Pro Patient standen z.B. die im Vorfeld einer anti-VEGF-Behandlung manuell erfassten Indikationen zur Verfügung. Visus (Sehschärfe) und Tensio (Augeninnendruck) wurden teilweise von den erfassenden Systemen in eine Datenbank eingespeist, wobei jedoch ein beträchtlicher Teil dieser beiden Parameter, ebenso wie Diagnosen nicht strukturiert erfasst war. Grund waren notfallmäßige Ambulanzbesuche, bei denen Visus und Tensio nicht elektronisch, sondern nur textuell erfasst wurden. Dasselbe gilt für alle Daten aus den ersten Jahren des Beobachtungszeitraums. Um zu allen Patienten vollständige, strukturierte und auswertbare Informationen für Auswertungen und Vorhersagen zu haben, sollten deshalb auch die numerischen Visus- und Tensio-Werte aus Arztbriefen extrahiert werden.

Informationen, die sowohl strukturiert als auch textuell vorliegen, eignen sich dann auch, um die Text-Mining-Extrakte anhand der strukturierten Daten zu evaluieren.

Tabelle 1 - Verfügbare Daten zum untersuchten Patientenkollektiv (3.683 Patientinnen und Patienten mit AMD, DMÖ oder RVO)

Art der Daten	Anzahl Datensätze
Visus	86.491
Tensio	53.720
Diagnosen	40.387
Arztbriefe	40.485

Die Datenanonymisierung war aufgrund einheitlicher Textmarkierungen in gleichartigen Arztbriefen problemlos. Das gesamte Datenmaterial ist demnach ohne Patientenbezug.

Grundsätzlich liegen zwei Arten von Arztbriefen vor: 1.) eine tabellarisches Textdokument, das bei jeder Sprechstundenvorstellung angelegt wird und 2.) eine Kurzversion davon, wenn der Patient zur weniger zeitaufwändigen OCT-Kontrolle bei ruhigem Netzhautbefund vorstellig wird (siehe Abb. 1).



UNIVERSITÄTSKLINIKUM FREIBURG
Klinik für Augenheilkunde Killianstr. 5, 79106 Freiburg

Klinik für Augenheilkunde, Universitätsklinikum Freiburg

Ärztlicher Direktor
Universitätsprofessor
Dr. Thomas Reinhard

Ambulanz
Freiburg, den

Killianstraße 5
79106 Freiburg



Sehr geehrte Frau Kollegin, sehr geehrter Herr Kollege,

wir berichten Ihnen nachfolgend über o.g. Patientin, die sich am 30.09.2015 in unserer OCT-Sprechstunde vorstellte.

Visus: Fernvisus 5m (Autorefraktometer)
RA: 0.4 (-1.0 / -2.5 / 80)
LA: 0.63 (-2.25 / -2.25 / 106)

OCT:

RA: neues intra- sowie subretinales Ödem

LA: neue Zysten intraretinal temporal oberhalb der Fovea --> Termin NH-Sprechstunde und Angiographie --> IVOM auch LA

Procedere:

Wir planen für das rechte Auge 3 Injektionen mit Lucentis im Abstand von 4 Wochen. Die genauen Termine werden Frau  zugeschickt.

Mit freundlichen Grüßen



Ärztin



Leitender Oberarzt

Verteiler: an Patienten

UNIVERSITÄTSKLINIKUM FREIBURG · Anstalt des öffentlichen Rechts · Sitz Freiburg
Aufsichtsrat Vorsitzender: Ministerialdirigent Clemens Benz
Vorstand Leitender Ärztlicher Direktor: Prof. Dr. Dr. h. c. mult. Jörg Rüdiger Stewert (Vorsitz)
Stellvertretender Leitender Ärztlicher Direktor: Prof. Dr. Dr. Rainer Schmelzisen
Kaufmännischer Direktor: Dipl.-Verwaltungswirt Bernd Sahner
Dekan der Medizinischen Fakultät: Prof. Dr. Kerstin Kriegelstein
Pflegedirektor: Helmut Schiffer

Bank Sparkasse Freiburg-Nördlicher Breisgau
BLZ 680 501 01 / Konto-Nr. 2004 406
IBAN DE08 6805 0101 0002 0044 06 / BIC FRSPDE66
Internet www.uniklinik-freiburg.de

KTQ®
-
zertifiziertes
Qualitäts-
management

Abbildung 1 - Beispielhafter Arztbrief

Eine Terminologie für die Augenheilkunde

Ziel des Text Mining war der Erhalt eines möglichst vollständigen Bildes aller Patienten. Neben Allgemein- und Begleiterkrankungen sollten die Extrakte ophthalmologische Merkmale zu den verschiedenen Augenabschnitten enthalten.

Symptome, Befunde und Erkrankungen werden routinemäßig mit einem ICD-10-Code versehen, z.B. H35.3 für "Degeneration der Makula und des hinteren Poles". H35.3 steht allerdings für eine ganze Reihe ähnlicher Merkmale (Tabelle 2), zwischen denen ICD-10 nicht differenziert. Eine Differenzierung zwischen einer feuchten und einer trockenen AMD ist in ICD-10 erst seit kurzem möglich.

Tabelle 2 - Beispiel eines ophthalmologischen ICD-Codes (Quelle: ICD-10-GM, Version DIMDI 2019)

H35.3	Degeneration der Makula und des hinteren Poles
	Inkl.:
	Drusen (degenerativ) Makula
	Fältelung Makula
	Gefäßähnliche Streifen [Angioid streaks] Makula
	Loch Makula
	Zyste Makula
	Kuhnt-Junius-Degeneration
	Senile Makuladegeneration (atrophisch) (exsudativ)
	Toxische Makulaerkrankung

ICD-10 wird deshalb in der Ophthalmologie primär für Abrechnungszwecke eingesetzt, und nur ihre Verfeinerung durch deutschsprachige Alpha-IDs macht klinisch Sinn [05]. Für eine präzise Kodierung im internationalen Kontext wäre aber ein internationaler Terminologiestandard wie SNOMED CT [07] besser geeignet. SNOMED CT erlaubt nicht nur eine wesentlich präzisere Kodierung als ICD-10, sondern umfasst neben Krankheiten auch Prozeduren, Organismen, Substanzen, Körperregionen etc. SNOMED-CT-Konzepte sind zum Großteil logisch definiert und in multiplen Hierarchien angeordnet, was Recherche und Weiterverarbeitung unterstützt. Seit März 2020 ist SNOMED CT in Deutschland für Forschungsprojekte, später auch für den Versorgungsbereich verfügbar. Allerdings existiert eine offizielle deutsche Übersetzung bisher nicht. TOPOs hat deshalb deutschsprachige

Fachtermini auf bestehende SNOMED CT-Codes gemappt, die von Ophthalmologen aus Arztbriefen extrahiert wurden, wobei sich 124 Vorzugsterme und 656 Synonyme ergaben².

Das Text-Mining-System Health Discovery

Zur Merkmalsextraktion aus den Arztbriefen kam die Text-Mining-Software Averbis Health Discovery (AHD)³ der Averbis GmbH zum Einsatz [08]. Damit können Texte nach einer Reihe von Merkmalstypen wie Krankheiten, Symptomen, Verordnungen oder Befunden analysiert und durchsucht werden. Hierfür werden sowohl statistische als auch regelbasierte Verfahren eingesetzt. Letztere sind typisch für die Erkennung fachsprachlicher Ausdrücke durch Abgleich mit einem vorgegebenen Vokabular, in dem der Text buchstaben- oder wortweise mit dem Vokabular abgeglichen wird und dann auf Terminologie-Codes abgebildet wird. Mit Ruta [09] lassen sich zudem regelbasiert komplexere Muster in Texten erkennen und daraus auf Merkmale schließen, die dann Auswertungen zur Verfügung stehen.

Die Erfassung des sprachlichen Kontexts ist eine zentrale Voraussetzung für korrekte Extraktion. Negationen sind in medizinischen Texten häufig (“Kein Anhalt für...”, “...konnte nicht nachgewiesen werden”), ebenso wie unsichere Aussagen (“Verdacht auf...”). Ebenso sind zeitlichen Kontext (“Zustand nach...”, “Der Visus war zuletzt...”) und Patientenbezug (Ausschluss von Familienanamnese) zu beachten. In der Augenheilkunde spielt zudem der Seitenbezug eine Rolle, da sich die meisten Angaben auf je ein Auge beziehen.

In dieser Arbeit konnte Health Discovery, unterstützt durch die maßgeschneiderte Terminologie und anhand eines angepassten Regelwerks, Angaben zu Visus, Tensio und Befundmerkmalen extrahieren. Abbildung 2 zeigt ein beispielhaftes Text-Mining-Resultat.

² Die Terminologie steht auf den Seiten der Averbis GmbH kostenfrei zum Download zur Verfügung: <https://topos.averbis.de/>

³ <https://health-discovery.io>

SNOMED

Diagnosen

Medikament

Visus

Tensio

Sehr geehrter Herr Kollege, sehr geehrter Patient,

wir berichten Ihnen nachfolgend über die Untersuchung vom XX.XX.XXXX in unserer Ambulanz.

Diagnosen

RA Feuchte altersabhängige Makuladegeneration, Z.n. 5 x Avastin

BA Pseudophakie

LA beginnender Nachstar

Allgemein art. Hypertonie, Hypercholesterinämie, ASS-Therapie, Rollstuhl

Vorgeschichte

Wiedervorstellung nach 2 weiteren Avastininjektionen im vergrößerten Abstand von 8 Wochen am rechten Auge. Das Sehen sei schwankend. Der Visus zuletzt war rechts cc 0,6 und links cc 1,0

Visus

RA 0,5 (eigene Brille)

LA 0,8 (eigene Brille)

Tensio

RA 14 mmHg

LA 11 mmHg

VAA

BA Deutliche Dermatochalasis, wenig Blepharitiszeichen, IOL in loco, LA leichter, fibrotischer Nachstar

Fundus

RA Papille schüsselförmig exkaviert, wenig Konus, Makula mit zentral pigmentierter Narbe, umgeben von Depigmentierung, keine Blutungen, keine Exsudate

LA Papille vital, guter Nervenfaserrandsaum, Makula mit PE-Verschiebungen

OCT

RA Zentrale Narbenbildung, nasal winzige neurosensorische Abhebung

LA Minimale PE-Verschiebungen, trocken

Beurteilung

Die heutige Untersuchung zeigt rechts nach vergrößertem Injektionsabstand von 8 Wochen einen stabilen Befund mit Narbenbildung und noch wenig subretinaler Flüssigkeit. Wir empfehlen keine weiteren Injektionen aber planen eine OCT-Kontrolle in 4 Wochen.

Mit freundlichen Grüßen

Abbildung 2 - Typisches Text-Mining-Ergebnis für einen ophthalmologischen Ambulanzbrief. Die deutschen SNOMED-CT-Übersetzungen sind hellgrün, Diagnosen nach ICD-10 gelb und Medikamente grau dargestellt. Die Visus-Werte sind dunkelgrün, die Tensio-Werte violett unterlegt.

Ergebnisse

Die Qualität des Text-Mining wurde für die Merkmale Visus, Tensio sowie die Diagnosen *feuchte AMD*, *DMÖ* und *RVO* ermittelt, indem vorhandene strukturierte Daten als Goldstandard verwendet wurden. Dabei wurden die Text-Mining-Ergebnisse mit denjenigen strukturierten Datenbankeinträgen verglichen, die im Behandlungsdatum übereinstimmten.

Insgesamt war in 47.535 Fällen ein Visuswert strukturiert vorhanden und zudem am selben Tag im Arztbrief dokumentiert (Tabelle 3, Spalte "Visus"). Für die Tensio bestand eine solche Doppeldokumentation in 40.358 Fällen (Spalte "Tensio").

Für die strukturiert erfassten Diagnosen waren Datumsangaben nicht verfügbar. Daher wurde pro Auge lediglich ermittelt, wie oft sowohl eine strukturiert erfasste als auch eine textbasierte Diagnose vorlag. Dies war für 3.211 Augen der Fall (Tabelle 3, Spalte "Diagnosen").

Die Ergebnisse der Untersuchung, wie oft ein Merkmal ausschließlich in Texten vorkam, also ohne Äquivalent in den strukturierten Datenbankeinträgen, finden sich in Tabelle 5 in der Zeile "Nur Text". Insgesamt fanden sich 18.932 Visus-Werte (+40%), 16.875 Tensio-Werte (+42%) sowie 4.103 Diagnosen (+128%) nur in Texten (Tabelle 3).

Tabelle 3: Verfügbarkeit strukturierter Daten im Vergleich zu allen Daten

Art der Datenerfassung	Visus		Tensio		Diagnosen	
	Anzahl	%	Anzahl	%	Anzahl	%
Strukturiert und Text	47.535		40.358		3.211	
Nur Text	18.932	+ 40%	16.875	+ 42%	4.103	+128%

Die sowohl strukturiert als auch textuell vorliegenden Merkmale wurden dann verwendet, um die Güte der Text-Mining-Verfahren in Form von Precision, Recall und F1-Score zu messen [10].

Dabei wurde vereinfachend davon ausgegangen, dass die "richtigen" Daten und somit der Goldstandard direkt vom messenden Gerät in der Datenbank abgelegt werden, was Übertragungsfehler ausschließen sollte. Diese Annahme wird jedoch durch die Fehleranalyse (siehe unten) relativiert.

Die Ergebnisse der Text-Mining-Evaluation anhand der strukturiert erfassten Werte finden sich in Tabelle 4. 97,6% der extrahierten Visus-Werte, 99,5% der extrahierten Tensio-Werte sowie 94,5% der extrahierten Diagnosen stimmten mit den strukturierten Datenbankeinträgen überein. Der Recall war insgesamt etwas niedriger: 91,9% der strukturiert erfassten Visus-

Werte, 97,1% der strukturiert erfassten Tensio-Werte und 95,0% der strukturiert erfassten Diagnosen wurde auch mittels Text-Mining erfasst. Daraus errechnete sich ein F1-Score von 94,7% für Visus und Diagnosen sowie 98,3% für die Tensio.

Tabelle 4: Übereinstimmung der Text-Mining-Ergebnisse mit den strukturierten Datenbankeinträgen

	Visus	Tensio	Diagnosen
Precision	97,6 %	99,5 %	94,5 %
Recall	91,9 %	97,1 %	95,0 %
F1-Wert	94,7 %	98,3 %	94,7 %

Fehleranalyse

Visus

Für eine Fehleranalyse bei der Visuswert-Extraktion wurden 100 zufällig ausgewählte Abweichungen aus dem oben genannten Testdatensatz ausgesucht und durch Experten überprüft. Die Abweichungen wurden in drei Kategorien eingruppiert (Tabelle 5).

Tabelle 5: Fehleranalyse der Visuswert-Extraktion mittels Text Mining. Ausgewertet wurden dabei die Fälle, in denen das Text Mining abweichende Ergebnisse von den Datenbankeinträgen der strukturierten Daten lieferte (5,3% der Gesamtheit aller Fälle)

Kategorie	Name	Häufigkeit	Beschreibung
1	Kein Visuswert im Text vorhanden	17 %	Kein Fehler des Text Mining, da im Arztbrief nur verkürzte Informationen ohne Visusangabe vorhanden waren.
2	Visuswert korrekt erkannt	22 %	Kein Fehler des Text Mining, da Visuswerte korrekt aus dem Text extrahiert wurden, jedoch nicht mit den entsprechenden Datenbankeinträgen übereinstimmten. Zum Teil ist dies bedingt dadurch, dass aus den Arztbriefen der <i>korrigierte</i> Visus (z.B. nach subjektivem Abgleich) extrahiert wurde. In der Datenbank war jedoch der Visus nur mit Autorefraktion eingetragen.
3	Visuswert nicht oder falsch erkannt	61 %	Entweder kein Visuswert erkannt oder fehlerhaftes Extrakt. Z.B. bedingt dadurch, dass eine Visus-Dokumentation von den ansonsten verwendeten Mustern stark abwich, oder dass ein Visus aus einer früheren Messung irrtümlich als aktueller Visus angesehen wurde.

Interessiert man sich lediglich für die Extraktionsgüte des Text-Mining und lässt für die Bewertung des Text-Mining-Algorithmus die Fehlerkategorien 1 und 2 der obigen Tabelle außer Betracht, da diese nicht auf Fehler des Text-Mining zurückzuführen sind, so ergibt sich

eine 39%ige Ergebnisverbesserung, oder in prozentualen Zahlen ausgedrückt eine Verbesserung des F1-Scores von 94,7 auf 96,5% (Tabelle 6).

Tabelle 6: Adjustierung der Güte der Visuswerte-Extraktion nach Fehleranalyse.

	Visus nicht adjustiert	Visus adjustiert
Precision	97,6 %	98,1 %
Recall	91,9 %	94,9 %
F1-Wert	94,7 %	96,5 %

Tensio

Die Fehleranalyse für die Tensio, ebenso wie die Kategorisierung erfolgte analog zum Visus auf 100 zufällig ausgewählten Abweichungen (Tabelle 7).

Tabelle 7: Fehleranalyse der Tensio-Extraktion mittels Text Mining. Ausgewertet wurden dabei die Fälle, in denen das Text Mining abweichende Ergebnisse von den Datenbankeinträgen der strukturierten Daten lieferte (1,7% der Gesamtheit aller Fälle)

Kategorie	Name	Häufigkeit	Beschreibung
1	Kein Tensiowert im Text	46 %	Kein Fehler des Text Mining, vielmehr enthielt der Arztbrief nur verkürzte Informationen ohne Tensio-Angabe.
2	Tensiowert korrekt erkannt	17 %	Kein Fehler des Text Mining, da der Tensiowert korrekt aus dem Text extrahiert wurde, wobei hier ein anderer Wert als in der Datenbank stand. Teilweise ist dies darauf zurückzuführen, dass manuell nachgemessen und der Wert nur im Brief korrigiert wurde.
3	Tensiowert nicht oder falsch erkannt	37 %	Entweder kein Tensiowert erkannt oder fehlerhaftes Extrakt. Teilweise ist dies darauf zurückzuführen, dass eine Tensio-Dokumentation von den sonst verwendeten Textmustern stark abwich, oder dass ein Tensio-Wert aus einer früheren Messung irrtümlich als aktueller Wert angesehen wurde.

Beschränkt man sich auf die Extraktionsgüte des Text-Mining und lässt die Fehlerkategorien 1 und 2 außer Betracht, resultiert eine Verbesserung der Tensiowert-Extraktion um 63%, und eine Verbesserung des F1-Scores von 98,3 auf 99,2% (Tabelle 8).

Tabelle 8: Adjustierung der Güte der Tensiowert-Extraktion nach Fehleranalyse.

	Tensio nicht adjustiert	Tensio adjustiert
Precision	99,5 %	99,6 %
Recall	97,1 %	98,9 %
F1-Wert	98,3 %	99,2 %

Diagnosen

Für die Diagnosen “feuchte AMD”, “RVO” und “DMÖ” wurde ebenfalls eine Fehleranalyse auf 100 zufällig ausgewählten Abweichungen aus demselben Testdatensatz durchgeführt. Die Abweichungen wurden in die gleichen drei Kategorien eingruppiert wie für Visus und Tensio (Tabelle 9).

Tabelle 9: Fehleranalyse der Diagnose-Extraktion mittels Text Mining. Ausgewertet wurden dabei die Fälle, in denen das Text Mining abweichende Ergebnisse von den Datenbankeinträgen der strukturierten Daten lieferte (5,3% der Gesamtheit aller Fälle)

Kategorie	Name	Häufigkeit	Beschreibung
1	Keine Diagnose im Text	6 %	Kein Fehler des Text Mining, da im Arztbrief nur eine verkürzte Informationen ohne Diagnoseangabe eingepflegt war.
2	Diagnose korrekt erkannt	44 %	Kein Text-Mining-Fehler, da die die Diagnose im Text korrekt extrahiert wurde, aber nicht mit der Diagnose in der Datenbank übereinstimmte. Teilweise ist dies darauf zurückführbar, dass beispielsweise die Diagnose “AMD” im Brief differenzierter aufgelöst wurde, als dies in der Datenbank der Fall war.
3	Diagnose nicht oder falsch erkannt	50%	Hier wurde eine im Text genannte Diagnose nicht erkannt, z.B. durch Schreibfehler oder ungewöhnliche Schreibvarianten. Auch kam vor, dass die Diagnose nicht direkt genannt, aber durch Kombination verschiedener Merkmale von einem Experten unmittelbar erschlossen werden kann.

Interessiert man sich wieder nur für die Extraktionsgüte des Text-Mining und lässt daher die Fehlerkategorien 1 und 2 außer Betracht, ergibt sich also nochmals eine 50%ige Ergebnisverbesserung, entsprechend einer Verbesserung des F1-Scores von 94,7 auf 97,1% (Tabelle 10).

Tabelle 10: Adjustierung der Güte der Diagnosenextraktion nach Fehleranalyse.

	Diagnose nicht adjustiert	Diagnose adjustiert
Precision	94,5 %	96,8 %
Recall	95,0 %	97,4 %
F1-Wert	94,7 %	97,1 %

Implementierung Kohortenfinder

Die Text-Mining-Ergebnisse dokumentierten eine hohe Zuverlässigkeit unseres Verfahrens, sodass wir es experimentell in einem Kohortenfinder implementierten. Dieser erlaubt aus den automatisch extrahierten Textmerkmalen hinreichend genaue Schätzungen von Patientenzahlen und -profilen, z.B. für Studienplanung und Prävalenzschätzungen. Folgende Fragestellung soll dies demonstrieren: "Wie viele Patienten mit neu diagnostizierter (*treatment naïver*) feuchter Makuladegeneration haben sich pro Monat in den vergangenen Jahren in Ihrer Ambulanz vorgestellt?"

Abb. 3 zeigt eine Übersicht, wie der Kohortenfinder die Anzahl ($n = 4.119$) der behandelten Augen für die drei Krankheiten "feuchte AMD", "DMÖ" und "RVO" darstellt. Für die Diagnose AMD sind im Betrachtungszeitraum 2.620 Augen, für RVO 762 und für DMÖ 733 Augen verfügbar. Der untere Teil der Übersicht zeigt zusätzlich die Verteilung der Diagnosen über die Jahre.

Anzahl Augen
4.119

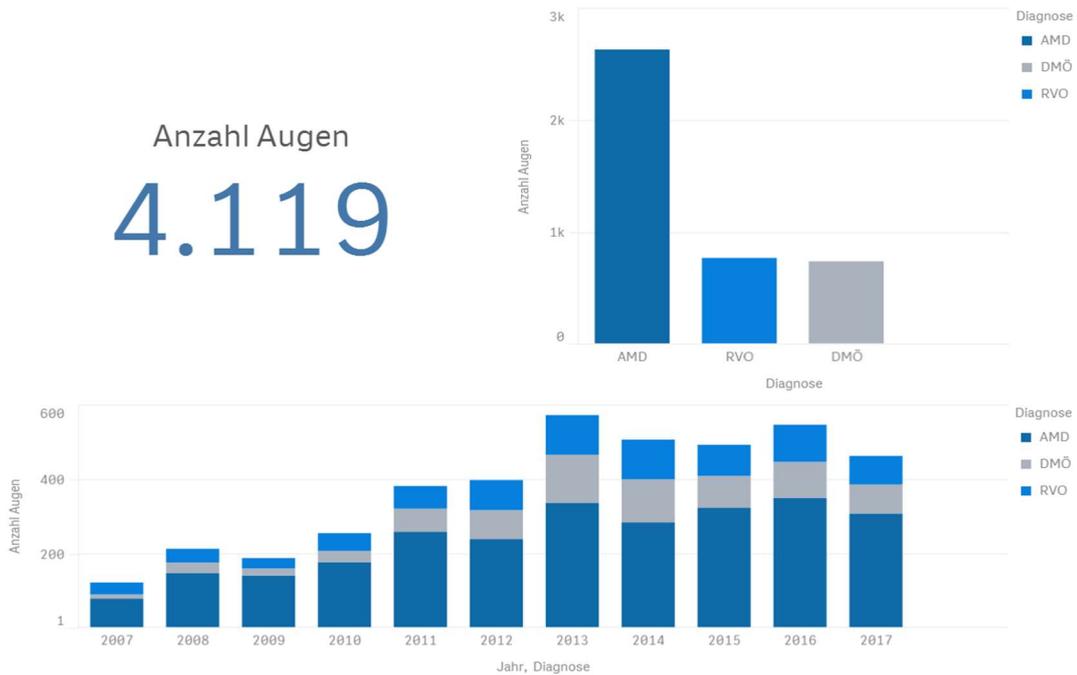


Abbildung 3 - Übersicht über die Anzahl der Patienten (Augen) und die Verteilung der Diagnosen "Feuchte AMD, RVO und DMÖ" insgesamt und über die Zeit.

Der Kohortenfinder ist interaktiv, so dass bei Klick auf einen Jahresbalken die Verteilung der drei Krankheiten über die einzelnen Monate des gewählten Jahres erscheint. Im Mittel sind dies monatlich etwa 30 Patienten mit AMD und je 8-9 Patienten für RVO und DMÖ. Darüber hinaus kann der Kohortenfinder auch Alters- und Geschlechterverteilung (Abbildungen 4 und 5) für die jeweiligen Diagnosegruppen darstellen.

AMD RVO DMÖ
78,7 70,6 66,4

Abbildung 4 - Übersicht über die Altersverteilung der Patienten bei erstmaliger Nennung der Diagnose in einem Arztbrief.

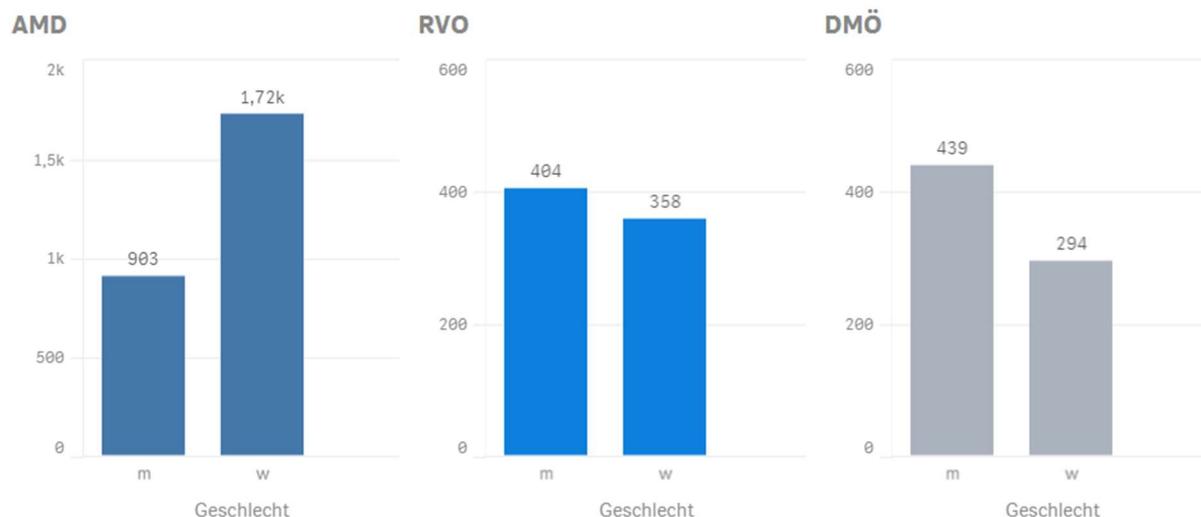


Abbildung 5 - Geschlechterverteilung für die drei Diagnosen "Feuchte AMD", "RVO" und "DMÖ"

Neben den hier genannten Beispielen lassen sich im Kohortenfinder viele weitere Visualisierungen individuell konfigurieren.

Analysetools wie der TOPOS-Kohortenfinder lassen prinzipiell mit gängigen Softwarewerkzeugen innerhalb moderner klinischer Arbeitsplatzsysteme oder Praxisinformationssystemen realisieren. Averbis Health Discovery wird hierbei über eine REST-Schnittstelle eingebunden⁴.

Diskussion

Anti-VEGF-Medikamente haben die Therapie der exsudativen Makuladegeneration sowie des Makulaödems bei Diabetes oder nach Venenverschluss revolutioniert. Entsprechend ist die Anzahl der durchgeführten intravitrealen Gaben von VEGF-Inhibitoren seit deren Einführung stetig angestiegen. 2017 lassen sich aus der Umfrage von DGII, BVA, BDOC und DOG 430.636 intravitreale Injektionen in Deutschland erfassen, was mehr als 1.000 Behandlungen täglich entspricht [11]. Insbesondere hinsichtlich demographischer Entwicklungen wird die Nachfrage weiterhin zunehmen [12]. Neben den eigentlichen Injektionen ist dabei zu beachten, dass jeder Therapiebesuch meist mehrere Kontrollbesuche beim Augenarzt vor und während der anti-VEGF-Therapie erfordert. Die hohe Zahl zusätzlicher Patientenkontakte stellt eine logistische Herausforderung dar, führt aber auch zu einer Fülle an Daten, die ausgewertet werden könnten, um Behandlungsmuster und -verläufe besser zu verstehen und

⁴ Auf Wunsch kann ein Zugang zu einer webbasierten Averbis Health Discovery - Instanz eingerichtet werden, worin dann die TOPOS-Extraktionspipeline getestet werden kann

möglicherweise auch vorherzusagen. Leider liegt dieser Schatz an klinischen Daten meist verborgen in Arztbriefen oder Ambulanzeinträgen, die einer systematischen Auswertung nicht leicht zugänglich sind. Ein Ziel des TOPOs-Projekts war die verlässliche Extraktion solcher Daten aus Arztbriefen mittels Text Mining, um darauf aufbauend Prognosemodelle für Therapieverläufe zu erstellen.

Bekannte Schwierigkeiten bei der Anwendung solcher Verfahren sind die korrekte Interpretation von Lateralität, zeitlichem Kontext, Negationen, Wahrscheinlichkeitsaussagen angesichts zahlreicher Schreibvarianten und -fehler. Damit muss ein Text-Mining-Verfahren umgehen können. In TOPOs hatten wir das Glück, über Werte zu Visus und Tensio sowohl strukturiert in Datenbanken als auch unstrukturiert in Arztbriefen zu verfügen. Damit konnte eine Messlatte an das Text-Mining-Verfahren angelegt werden, wobei es sich bei der Extraktion von Visus- und Tensio-Werten sowie von Diagnosen als präzises Werkzeug erwiesen hat. Text Mining kann auch die Identifikation von Studienkohorten unterstützen. So kann ermittelt werden, wie viele Patienten mit einer bestimmten Diagnose im Gesamtkollektiv einer Praxis oder Klinik für eine geplante klinische Studie verfügbar sind [13].

Die Entscheidung, ophthalmologische Fachtermini auf SNOMED-CT [07][14] abzubilden, erforderte die manuelle Erstellung eines maßgeschneiderten Vokabulars⁵. Mit einer Erweiterung solcher Sammlungen fachsprachlicher Ausdrücke und deren Abbildung auf einen Terminologiestandard [15] kann diese Anwendung auch über die Augenheilkunde hinaus, z.B. zur Diagnosenunterstützung erweitert werden, wie für den Anwendungsfall seltene Erkrankungen bereits demonstriert [16].

Fazit für die Praxis

Nachdem sich die intravitreale anti-VEGF-Therapie vor allem bei der exsudativen altersabhängigen Makuladegeneration, dem diabetischen Makulaödem sowie dem Makulaödem nach Gefäßverschluss seit Jahren als Goldstandard etabliert hat, liegen nun an vielen Zentren große Datenmengen über nennenswerte Zeiträume vor, die zur Beantwortung von epidemiologischen und wissenschaftlichen Fragestellungen von höchstem Wert sind. Bisher verhinderte die unstrukturierte Beschaffenheit dieser Daten häufig eine direkte Nutzung, da sie im Behandlungsalltag und nicht unter Studienbedingungen erhoben und damit meist nicht standardisiert erfasst wurden. Unser Ansatz zur Datenanalyse mittels Text-Mining ermöglicht die Nutzung der Behandlungsdaten und eröffnet eine neue Informationsquelle für wissenschaftliche Fragestellungen. Interessante zukünftige Fragestellungen wären

⁵ Über die Averbis-Homepage kostenlos verfügbar

beispielsweise die Auswertung größerer Datenmengen aus heterogenen Datenquellen, beispielsweise im Rahmen von Registerprojekten und unter Einbeziehung unterschiedlicher Praxis- oder Klinik-EDV-Systeme.

Danksagungen

TOPOs wurde vom BMBF (Bundesministerium für Bildung und Forschung) im Rahmen des Förderschwerpunktes „Digitale Gesundheitsversorgung“ gefördert.

Referenzen

- [01] Griffon N, Charlet J, Darmoni SJ. Managing free text for secondary use of health data. *Yearb Med Inform.* 2014 Aug 15;9:167-169.
- [02] Schlegel DR, Ficheur G. Secondary Use of Patient Data: Review of the Literature Published in 2016. *Yearb Med Inform.* 2017 08;26(1):68-71.
- [03] Safran C, Bloomrosen M, Hammond WE, Labkoff S, Markel-Fox S, Tang PC, et al. Toward a national framework for the secondary use of health data: an American Medical Informatics Association White Paper. *J Am Med Inform Assoc.* 2007 Jan-Feb;14(1):1-9.
- [04] Wang Y, Wang L, Rastegar-Mojarad M, Moon S, Shen F, Afzal N, et al. Clinical information extraction applications: A literature review. *J Biomed Inform.* 2018 01;77:34-49.
- [05] Böhringer D, Lang SJ, Daniel MC, Lapp T, Reinhard T. Automatisierte Zuordnung von ICD- und Alpha-ID-Codes zu „Real-World“-Arztbriefdiagnosen durch die „word2vec“-Methode. *Klin Monbl Augenheilkd.* 2019 Dec;236(12):1413-1417.
- [06] Névéol A, Dalianis H, Velupillai S, Savova G, Zweigenbaum P. Clinical natural language processing in languages other than English: opportunities and challenges. *J Biomed Semantics.* 2018 03 30;9(1):12.
- [07] Bodenreider O, Cornet R, Vreeman DJ. Recent Developments in Clinical Terminologies – SNOMED CT, LOINC, and RxNorm. *Yearb Med Inform.* 2018 Aug;27(1):129-139.
- [08] Pokora RM, Cornet LL, Daumke P, Mildenerger P, Zeeb H, Blettner M. Validierung von semantischen Analysen von unstrukturierten medizinischen Daten für Forschungszwecke. *Gesundheitswesen.* [Publiziert online: 9. Oktober 2019].
- [09] Kluegl P, Toepfer M, Beck PD, Fette G, Puppe F. UIMA Ruta: Rapid development of rule-based information extraction applications. *Natural Language Engineering,* 2016; 22(1): 1-40.

- [10] Beurteilung eines binären Klassifikators. Wikipedia. https://de.wikipedia.org/wiki/Beurteilung_eines_binären_Klassifikators. Letzter Zugriff 30.1.2020
- [11] Wenzel M, Dick HB, Scharrer A, Schayan K, Reinhard T. Umfrage von BDOC, BVA, DGII und DOG zur ambulanten und stationären Intraokularchirurgie: Ergebnisse für das Jahr 2017. Ophthalmo-Chirurgie 2018; 30: 255-266
- [12] Li JQ, Welchowski T, Schmid M, Mauschwitz MM, Holz FG, Finger RP. Prevalence and incidence of age-related macular degeneration in Europe: a systematic review and meta-analysis. Br J Ophthalmol. [publiziert online: 11. November 2019].
- [13] O'Mara-Eves A, Thomas J, McNaught J, Miwa M, Ananiadou S. Using text mining for study identification in systematic reviews: a systematic review of current approaches. Syst Rev. 2015 Jan 14;4:5.
- [14] Willett DL, Kannan V, Chu L, Buchanan JR, Velasco FT, Clark JD, et al. SNOMED CT concept hierarchies for sharing definitions of clinical conditions using electronic health record data. Appl Clin Inform. 2018 07;9(3):667-682.
- [15] Hashemian Nik D, Kasáč Z, Goda Z, Semlitsch A, Schulz S. Building an experimental German user interface terminology linked to SNOMED CT. Stud Health Technol Inform. 2019 Aug 21;264:153-157.
- [16] Shen F, Zhao Y, Wang L, Mojarad MR, Wang Y, Liu S, et al. Rare disease knowledge enrichment through a data-driven approach. BMC Med Inform Decis Mak. 2019 02 14;19(1):32.

Interessenkonflikte

Die Autoren Bastian Grundel, Marc-Antoine Bernardeau, Holger Langner, Christoph Schmidt, Daniel Böhringer, Marc Ritter, Paul Rosenthal und Andreas Stahl geben an, dass kein Interessenkonflikt besteht.

Andrea Grandjean und Stefan Schulz stehen in einem Beschäftigungsverhältnis zur Fa. Averbis GmbH, die das in der Studie verwendete System Health Discovery entwickelt und vertreibt. Philipp Daumke ist geschäftsführender Gesellschafter der Averbis GmbH.

Patientenrechte

Für die Studie kamen ausschließlich anonymisierte Klinikdaten zur Anwendung, so dass nach geltendem Recht kein Votum einer Ethikkommission erforderlich war. Für die Vollständigkeit der manuell durchgeführten Anonymisierung verbürgen sich die klinischen Projektpartner.