FAKULTÄT FÜR !NFORMATIK
Faculty of Informatics

Electronic Health Record Systems in General Practice

Video Analysis of Medical Consultations and Implications for Secondary Analysis

DIPLOMARBEIT

zur Erlangung des akademischen Grades

Diplom-Ingenieur

im Rahmen des Studiums

Medieninformatik

eingereicht von

Christian Löw, B.Sc.
Matrikelnummer 0425873

an der Fakultät für Informatik
der Technischen Universität Wien
Betreuung: Univ.Prof. Geraldine Fitzpatrick, Ph.D.

Wien, 24. Februar 2016

Christian Löw	Geraldine Fitzpatrick

FAKULTÄT FÜR !NFORMATIK
Faculty of Informatics

Electronic Health Record Systems in General Practice

Video Analysis of Medical Consultations and Implications for Secondary Analysis

DIPLOMA THESIS

submitted in partial fulfillment of the requirements for the degree of

Diplom-Ingenieur

in

Media Informatics

by

Christian Löw, B.Sc.
Registration Number 0425873

to the Faculty of Informatics

at the Vienna University of Technology

Advisor: Univ.Prof. Geraldine Fitzpatrick, Ph.D.

Vienna, 24th February, 2016

Christian Löw

Geraldine Fitzpatrick

Erklärung zur Verfassung der Arbeit

Christian Löw, B.Sc.
Garnisongasse 24/14, A-1090 Wien.

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

Wien, 24. Februar 2016

Christian Löw

Abstract

The design of Electronic Health Record (EHR) systems remains a challenging task, as health care systems, professions and institutions in themselves and at the interfaces of their collaboration yield a complex, intricate concept of (cooperative) work in a professional context at the intersection of systematization and care-giving. EHR documentation data furthermore is not only used in settings of primary cooperative work, but also taken for secondary analysis; a broadly used tool, used by, among others, epidemiologists to better understand disease spread, prevalence or morbidity by conducting statistical analyses. Such secondary reasoning makes assumptions about the EHR data at hand, for example in terms of its content and validity: It is often taken at face value, neglecting its situatedness.

This thesis builds on a field study conducted in the United Kingdom on the work practice of general practitioners (GPs), in which course GPs were recorded in their work with patient-actors, who enacted two pre-defined scenarios. In a first step, it is the purpose of this work to obtain an in-depth understanding of how these scenarios are treated and subsequently documented in the EHR system by the GPs. To this end, video recordings and EHR system screen recordings of 30 consultations, conducted by 15 GPs of five general practice surgeries as well as walkthrough interviews with the attending GPs were subjected to ethnographic analysis, yielding a variety of variabilities and influencing factors in case documentation across surgeries, GPs and the three different EHR system brands at hand (Vision, EMIS LV, SystmOne).

In a further step, these results are put in relation to aspects of secondary analysis. To this end, interviews were conducted with experts in the fields of epidemiology, Natural Language Processing (NLP), and EHR system design, relating the analysis findings to prevailing concepts of data validity and content extraction. This revealed a rich picture of the practice of the GPs' work and needs directed at EHR systems as well as means of structuring EHR documentation on one hand and incentives to formalize their EHR documentation directed at them from the „outside world“ that is secondary analysis to it on the other hand. The thesis concludes with recommendations directed at the design of EHR systems and EHR data structures that aim at reducing variability in EHR documentation.

Kurzfassung

Die Gestaltung von Electronic Health Care (EHR)-Systemen stellt eine herausfordernde Aufgabe dar, da die Systeme, Professionen und Institutionen des Gesundheitswesens an der Schnittstelle ihrer Zusammenarbeit ein komplexes, vielfältig verwobenes Konzept von (kooperativer) Arbeit in einem gemeinsamen professionellen Kontext an der Ligatur von Systematisierung und medizinischer Fürsorge verbindet. EHR-Dokumentation findet darüber hinaus nicht nur in Zusammenhängen primärer Kooperativarbeit Verwendung, sondern wird auch für Sekundäranalyse herangezogen; ein vielfältig genutztes Werkzeug, das, unter anderem, von Epidemiologen zum besseren Verständnis von Verbreitung und Häufigkeit von Krankheiten sowie zur Durchführung statistischer Analysen verwendet wird. Solche Formen weiterführender Verarbeitung stellen Vermutungen über die genutzten Daten an, zum Beispiel bezüglich ihres Inhalts und ihrer Verlässlichkeit: Sie werden oft wenig hinterfragt verwendet, wodurch ihre Situiertheit vernachlässigt bleibt.

Diese Diplomarbeit baut auf einer im Vereinigten Königreich durchgeführten Feldstudie über die Arbeitspraxis von AllgemeinmedizinerInnen auf, in deren Zuge diese in ihrer Arbeit mit PatientendarstellerInnen, die zwei vordefinierte Szenarien vorspielten, videoaufgezeichnet wurden. In einem ersten Schritt verfolgt die vorliegende Arbeit den Zweck, ein tiefgreifendes Verständnis darüber zu entwickeln, wie diese Szenarien behandelt und weiters von den AllgemeinmedizinerInnen im EHR-System dokumentiert wurden. Dazu wurden Videoaufzeichnungen und EHR-System-Bildschirmmitschnitte von 30 Konsultationen, durchgeführt von 15 AllgemeinmedizinerInnen in fünf Arztpraxen, sowie Nachbesprechungsinterviews mit den behandelnden ÄrztInnen ethnografisch analysiert, was eine Vielzahl an Variabilitäten und beeinflussenden Faktoren in der Falldokumentation unter den Arztpraxen, AllgemeinmedizinerInnen sowie den drei verschiedenen, von den ÄrztInnen verwendeten EHR-Systemen (Vision, EMIS LV, SystmOne) ergab.

In einem weiteren Schritt werden diese Resultate vor dem Hintergrund von Aspekten der Sekundäranalyse betrachtet. Dazu wurden eine Literaturanalyse sowie semistrukturierte Interviews mit ExpertInnen in den Bereichen Epidemiologie, Natural Language Processing (NLP) und EHR-Systemgestaltung durchgeführt, die die Analyseergebnisse mit vorherrschenden Konzepten von Datenverlässlichkeit und Inhaltsextraktion in Beziehung setzen. Dies ergab ein reichhaltiges Bild von an die EHR-Systeme und darin verfügbare Möglichkeiten der Datenstrukturierung gerichteten Bedürfnissen der AllgemeinmedizinerInnen einerseits und von der Äussenwelt", die die Sekundäranalyse zur

Arbeitspraxis der Ärzte darstellt, an sie herangetragene Anreize, ihre Dokumentation darüber hinausgehend zu strukturieren, andererseits. Die Diplomarbeit formuliert an die Gestaltung von EHR-Systeme und EHR-Datenstrukturen gerichtete Empfehlungen, die darauf abzielen, die in EHR-Dokumentation bestehende Variabilität zu reduzieren.

Acknowledgements

First and foremost, I would like to thank my supervisor, Prof. Geraldine Fitzpatrick, for the opportunity to work on the basis of the PREP project, her encouragement throughout my engagement with the topic of this thesis work and generally the very supportive and curiosity-sparking environment provided in the course of this work as well as in her classes, which I always loved to attend.

Furthermore, I would like to thank the investigators involved with the PREP field study for providing the basis to a very interesting diploma project, as well as the experts interviewed in the course of it.

This thesis constitutes the end of a journey, and expressions of gratitude to a long list of fine and lovely people are in order: To Christoph and Thomas for the conversations and critical reflection of this thesis work. To many other colleagues whom I enjoyed studying with over the course of my life as a student at the TU Vienna. To my family - especially my parents and my siblings Valerie and Michael - for your unconditional support. To Laura, Christoph, Johannes, Dani, Christoph, Thomas, Stefan and many more for your friendship. To my fellow scouts for being who you are.

Contents

Abstract **vii**

Kurzfassung **ix**

Acronyms **1**

1 Introduction **3**
1.1 Problem Setting . . . 3
1.2 Attribution . . . 4
1.3 Subject of This Work . . . 4
1.4 Methodological Overview . . . 5
1.5 Thesis Structure Outline . . . 6

2 Background and Related Work **9**
2.1 EHR Systems in Health Care . . . 9
2.2 Secondary Analysis of Primary Care EHR Data . . . 10
2.3 The Patient Records Enhancement Programme (PREP): The Field Study 11
2.4 Structure in EHR Systems . . . 15

3 Methodology **17**
3.1 Paradigmatic Stance . . . 17
3.2 Methodological Considerations on Video Analysis . . . 18
3.3 Analysis of PREP Field Study Data . . . 22
3.4 Literature Review . . . 29
3.5 Expert Interviews . . . 30

4 Jane and Scott in Consultation **33**
4.1 Overview of Scenarios . . . 33
4.2 Lifelikeness and Consistency of Simulated Patients . . . 37
4.3 Basic Focal Themes in a Consultation . . . 39
4.4 Overview of Consultation Outcomes . . . 45
4.5 Summary and Conclusion . . . 46

5 Jane and Scott in EHR Documentation **51**

5.1 Comparing EMIS LV, SystmOne and Vision . . . 51
5.2 Content Analysis of EHR Documentation . . . 64
5.3 Summary and Conclusion . . . 78

6 Working with EHR Systems **81**
6.1 On Coding: *Codes Are Useful* . . . 82
6.2 On Free Text: An *Important* Narrative . . . 86
6.3 On Keeping Electronic Health Records . . . 88
6.4 Summary and Conclusion . . . 94

7 Secondary Analysis of Primary Care EHR Data **97**
7.1 Definition and Applications . . . 98
7.2 Data Sources for Secondary Analysis . . . 100
7.3 Dimensions of Data Quality and Strategies of Validation . . . 102
7.4 Phenomena of Variability . . . 108
7.5 Summary . . . 116

8 Discussion and Conclusion **117**
8.1 Overview of Results . . . 117
8.2 Contribution of this Work . . . 121
8.3 Implications for Design . . . 122
8.4 Limitations . . . 127
8.5 Conclusion and Future Work . . . 127

Bibliography **129**

A Appendix: Video Analysis Process **137**
A.1 Field Study Data Sources . . . 137
A.2 Example EHR Documentation of *Scott* and *Jane* case . . . 137

B Appendix: Expert Interviews **141**
B.1 Interview Outline - Epidemiologist . . . 141
B.2 Interview Transcript Example - Epidemiologist . . . 142
B.3 Consent Form . . . 149

C Appendix: Compilation of EHR Usage Charts **151**
C.1 Category Usage Schematics . . . 151
C.2 Word Counts . . . 151
C.3 Code and Category Usage . . . 151
C.4 Word Count and Age . . . 151

List of Figures **159**

List of Tables **162**

Acronyms

CPRD Clinical Practice Research Datalink.

CSCW Computer Supported Cooperative Work.

DIN Doctors' Independent Network.

EHR Electronic Health Record.

FPS Frame per Second.

GP General Practitioner.

GPRD General Primary Care Database.

HCI Human Computer Interaction.

HSR Health Services Research.

ICD International Classification of Diseases.

ICPC International Classification for Primary Care.

JP Joint Pain.

LKF Leistungsorientierte Krankenanstalten-Finanzierung.

MNAR Missing Not at Random.

NHS National Health Service.

NLP Natural Language Processing.

PREP Patient Records Enhancement Programme.

QOF Quality and Outcomes Framework.

SP Simulated Patient.

ST Sore Throat.

THIN The Health Improvement Network.

VIA Video Interaction Analysis.

WHO World Health Organization.

CHAPTER 1

Introduction

1.1 Problem Setting

Electronic Health Record (EHR) systems are employed in many different strands of health care, assisting involved professionals in documentation, reasoning and cooperative work. The design of good EHR systems remains a challenging task, as health care systems, professions and institutions in themselves and at the interfaces of their collaboration yield a complex, intricate concept of (cooperative) work in a professional context at the ligature of systematization and care-giving. IT systems hold opportunities, offer potential assistance and improvement of quality, but need to be assessed in their effects on the workplace they are supposed to assist in, as they attempt to control and shape workflows and interaction and by that also affect the underlying understanding of this work.

EHR documentation data furthermore is not only used in settings of „direct“ cooperative work, but also taken for subsequent, secondary reasoning. Such secondary analysis is a broadly used tool, e.g. used by epidemiologists to better understand disease spread, prevalence or morbidity by conducting descriptive, inferential or predictive statistical analyses which rely on and also make assumptions about the EHR data at hand, for example in regard of its content and reliability: It is often taken at face value, neglecting its situatedness. Data originating from general practitioner (GP) surgeries represents an important source for secondary analysis, that is e.g. aggregated in the Clinical Practice Research Datalink (CPRD) database .

This thesis builds on data gathered during a study that was conducted in the course of the Patients Record Enhancement Project (PREP), an interdisciplinary research program carried out by Brighton and Sussex Medical School, University of Sussex and University of Brighton (United Kingdom). PREP was a research project aimed at better understanding the effect of EHR systems in the context of primary care, where general practitioners usually are offered systematized code sets and free-text fields to describe a patient and their work with them. To this end, a field study was carried out: Using two

simulated patients, who enacted scenarios of sore throat and joint pain symptomatics, consultations with 15 general practitioners in five primary care surgeries were video-recorded. Furthermore, EHR system usage was recorded by screen capture software, and the involved general practitioners were interviewed following their consultation. Further information on the PREP field study can be found in chapter 2.2.

1.2 Attribution

The author of this thesis would like to thank the investigators of the PREP project who conceived and conducted the field study which forms the basis of this master's thesis work. This project is outlined in section 2.3.

1.3 Subject of This Work

One of the main subjects of this work was to analyze the video recordings and screen recordings of the consultations, the resulting EHR documentation and the walkthrough interviews conducted with the GPs. To guide this analysis, research questions (RQs) were put in place that aim at identifying possible forms of variability in consultations and EHR documentation across the three different brands of EHR system employed (Vision, SystmOne and EMIS LV) as well as across the surgeries:

- RQ 1: *How are the two simulated scenarios dealt with in the consultations and how is the information originating from the consultations documented in the EHR systems at hand?*
- RQ 2: *How does this relate to the understanding GPs have of the building blocks of EHR documentation - coded entries and free text?*

This to the end of describing usage practice in general as well as the effect possible forms of variability have on the resulting EHR data in order to better understand how and why EHR documentation comes to be.

In a further step, such usage practice and forms of variability are interpreted in respect to efforts of secondary analysis of primary care EHR data:

- RQ 3: *How is EHR data coming from general practice seen by epidemiologists conducting secondary analysis with this data?*

The overall goal of the work is to describe variability, interpret it in its implications for efforts of secondary analysis and draw design guidelines from this interplay towards possible improvement of EHR system design.

Results

This work aimed at developing a detailed picture of GP consultations and the usage of EHR systems. The walkthrough interviews provided background on the usage practices and opinions of the GPs in respect to coding and free text practices. The resulting picture of GP practice in EHR usage exposed variability across the three EHR systems used by the 15 GPs and five surgeries involved.

Implications on secondary analysis unfolded in terms of a gap of meaning between the GPs' intention in documenting their actions and diagnoses in the EHR system on one side and the assumptions and interpretations attributed to the resulting EHR data by secondary analysis researchers, e.g. in terms of completeness, on another side. Investigating the contrast between the pragmatic needs GP practice directs at the design of EHR systems on one hand and the assumptions about the resulting data made in secondary analysis on the other hand (see also figure 1.1) might generally contribute to a better understanding on how EHR data gets constructed, how and which meaning is packed into it and how it relates to the consultation it depicts.
Finally, this work yielded implications for design on how to better handle this field of tension in the design of EHR systems.

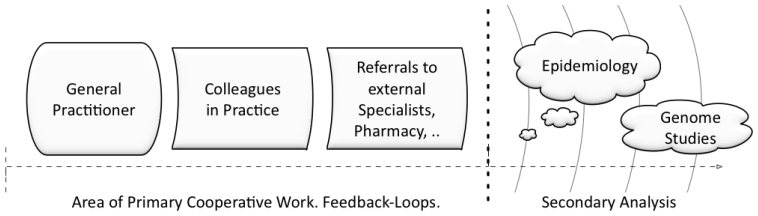


Figure 1.1: Usage of EHRs among primary actors (e.g. GP, colleagues, specialists, ..) is embedded in collaboration that, at least to some part, can rely on the „ground truth" of patient care to negotiate and validate its process. Secondary Analysis is disjunct from this system and often has to make assumptions on how the collaborative artefacts that constitute EHR documentation are to be interpreted.

1.4 Methodological Overview

The overall methodological approach of this thesis is set up under a constructivist paradigm and qualitative, ethnographic methodology. The usage practice of GPs with EHR systems are understood as a result of a continuous process of construction and evolution of the understanding of care-giving, patient-doctor interaction and organizational (sub)routines supported by EHR systems. As an approach to capturing cooperative work in GP surgeries, Swinglehurst et al. [1] refer to a model of organizational routines, that differentiates between the ostensive, an abstract idea of a work routine, the performative, its practical realization and possible artefacts codifying and distributing aspects of the routine.

Analysis of PREP Field Study Data

In the course of the work done in this thesis with the PREP field study's data, 23 hours of video recordings that have been collected during the PREP field study were analyzed. This encompasses (1) ca. 12 hours of consultations conducted by 15 general practitioners that work in five surgeries, (2) 4 hours of walkthrough interviews conducted with the GPs and (3) 7 hours of screen recordings of the EHR systems as they are used by the GPs.
To this end, multiple, iterative passes of viewing, annotative coding and analysis will be done. The video recordings of consultations will be approached by discourse analysis as done in related ethnographic case studies in primary care surgeries conducted by Swinglehurst et al. [3] [10] and Roberts and Sarangi [11].
The methodology of this video-analysis is informed by „video interaction analysis“ (VIA), a concept of the analysis of video-data proposed by Hubert Knoblauch [12], who appropriated conversation analysis, a concept originating from analysis on audio-recordings, to the means of video. In general, the analysis of the consultation recordings, the walkthrough interviews, screen recordings, EHR documentation and field notes will be embedded in a process of inductive theory generation as Grounded Theory. Furthermore, the EHR documentation created by the GPs was subjected to a content analysis. This was done in order to obtain a detailed idea on how the work practice as a GP translates into EHR documentation and what information it contains.

Implications for Secondary Analysis

To find implications of the field study analysis results for secondary analysis, a literature review was conducted on secondary analysis aspects and specifically on epidemiological publications of secondary analysis research done on basis of the primary care EHR data made available for research in the CPRD database. Furthermore, semi-structured expert interviews were conducted with experts relevant to EHR design and secondary analysis, focusing on EHR experts, NLP experts, as well as researchers involved in secondary analysis, e.g. in epidemiology. Several experts in these fields have already been contacted and declared to be available for interviewing. The resulting insights were related to the design of EHR systems in the form of design guidelines.

1.5 Thesis Structure Outline

Chapter 2: Background and Related Work describes the background of this work and related work. **Chapter 3: Methodology** describes the methodology employed in this thesis. **Chapter 4: *Jane* and *Scott* in Consultation** describes the video analysis results of the consultations with *Scott* and *Jane*, the two simulated patients. This includes reflections on the performance of the simulated patients, a description of the course of the consultations along common basic focal themes as well as a description of variability in GP approach and outcomes. **Chapter 5: *Jane* and *Scott* in EHR Documentation** describes how the simulated scenarios were documented in the EHR

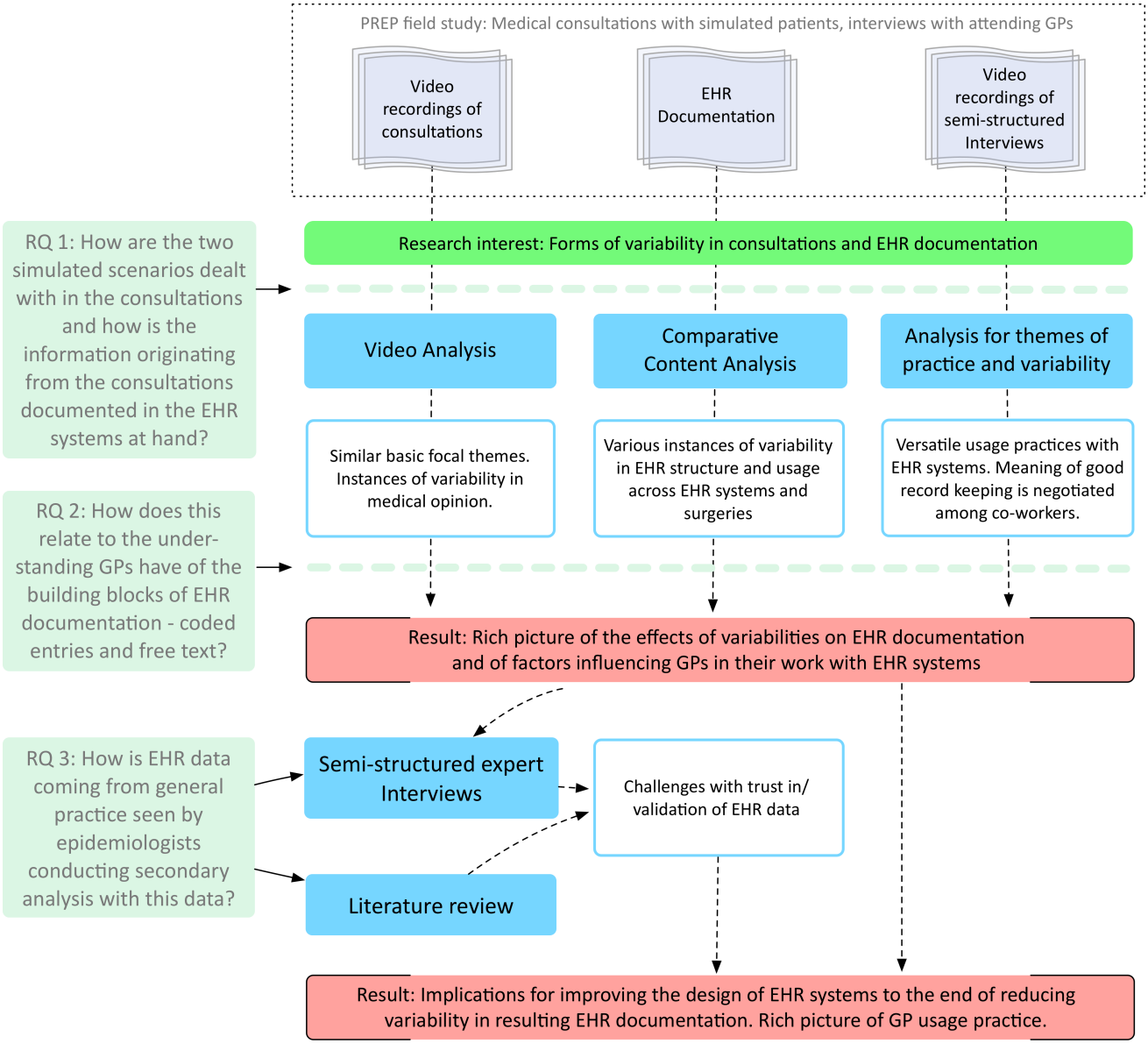


Figure 1.2: Outline of the several tasks and related results.

systems by means of comparative content analysis. After describing and comparing EMIS LV, SystmOne and Vision, the three different EHR systems employed by surgeries HS, EH, OS, MS and SE, results from an in-depth analysis of the EHR entries are described. This chapter concludes with various instances of variability across EHR systems and surgeries. **Chapter 6: Working with EHR Systems** describes themes of usage practice and variability that emerged from an analysis of the walkthrough interviews conducted with the attending GPs. **Chapter 7: Secondary Analysis of Primary Care EHR Data** describes the results of semi-structured interviews conducted with epidemiologists and other experts as well as a literature review, highlighting challenges with trusting EHR data when using it fur purposes of secondary analysis. **Chapter 8: Discussion and Conclusion** summarizes the themes of variability that emerged with the analysis of the PREP field study's data and, in light of the results of chapter 7, yields implications for improving the design of EHR systems towards the reconciliation of the needs of GPs directed at EHR systems on one hand and requirements for data-usage for purposes of secondary analysis on the other hand. The thesis concludes with suggestions for future work. **Appendix**: The appendix of this work contains description on the data sources as well as the EHR documentation for reference, as well as a compilation of the EHR documentation analysis charts used in chapters 5 and 6.

CHAPTER 2

Background and Related Work

This chapter describes the background and related work that served as a point of departure for this master's thesis. Section 2.1 (*EHR Systems in Health Care*) describes related ethnographic studies that were taken on as methodical references. Section 2.2 (*Secondary Analysis of Primary Care EHR Data*) describes aspects of secondary analysis that served as point of reference for formulating related implications. Section 2.3 (*The Patient Records Enhancement Programme: The Field Study*) outlines aspects of the PREP field study as the source of the data analyzed in this master's thesis. Section 2.4 (*Structure in EHR Systems*) describes background of structure in EHR systems.

2.1 EHR Systems in Health Care

The question of how the complex practice of cooperative work in health care can be supported by technology is of long-standing interest to efforts of effective, sustainable design of IT for health care. Supporting cooperative work in health care contexts demands for a through understanding of the practice to be supported, as described e.g. by Heath and Luff, who argued, that what might seem as 'bad' usage of EHR systems in an individual perspective (e.g. not abiding to defined content structure) relates to justified, important needs and cooperative work aspects hard to understand and support with simple data handling tools [2].
This thesis relates to several ethnographic inquiries conducted on the usage practice of Primary Care EHR systems.

- In four UK general practice surgeries, Swinglehurst et al. (2012a) conducted interviews, ethnographic observations as well as video recordings and screen recordings, collected various artifacts like documents in order to conduct a qualitative inquiry on the conceptualization and enactment of the work practice of General Practitioners (GPs) with special focus on prescription, coding and summarizing as well as

chronic disease surveillance. In front of the background of sociological theory, they concluded that observing collaborative work routines hold the potential to identify various systemic problems, like poor performance or failures of coordination) as well as "*reveal the hidden work and workarounds by front-line staff which bridge the model-reality gap in EPR technologies*" [3, p. 1].

- Further related work by Swinglehurst et al. (2012b) constitutes an ethnographic case study on the usage practice of electronic templates in chronic disease management focusing on "*how electronic templates shape, enable and constrain consultations about chronic diseases*" [1].
- Hardstone et al. conducted an ethnographic study on the work of mental health teams and the deployment of an EHR system, finding that the practice of EHR related record-keeping is influenced significantly by aspects of informality, and concepts of intended usage shaping the design of EHR systems are met in practice by a complex interplay of verbal communication, handwritten records and electronic records [4].

Common to these works and taken on as a perspective important to this master's thesis is a sociotechnical approach to EHR systems:

> We adopted a sociotechnical approach, meaning we focused on the dynamic, contingent interaction between humans and technologies rather than assuming technology is itself 'causal' of specific effects. [1, p. 2]

In a literature review on EHR research, Greenhalgh et al. concluded, that EHR systems will "*always require human input to recontextualize knowledge*", which further underlines the importance of a broad, social-science-informed perspective [5]; it is backed by various theoretical underpinnings explored in the works of Feldman and Pentland on the artefacts of organizational routines [6], and of Roberts and Sarangi on theme-oriented discourse analysis of medical consultations [7]; these aspects are described in more detail in chapter 3 (*Methodology*).

2.2 Secondary Analysis of Primary Care EHR Data

Using EHR data for secondary analysis is subject to research in terms of aspects of the validity of this data for such purposes as well as strategies of validation. The works described in this section will only be mentioned briefly, as they are discussed in more detail in chapter 7 (*Secondary Analysis of Primary Care EHR Data*).

- Boslaugh gave a comprehensive introduction to secondary analysis in the fields of epidemiology and public health, detailing on the necessity to acquire a detailed

understanding of the data to be taken on for secondary analysis and its situatedness in the circumstances it primarily originated from [8].

- Khan et al. conducted a review of research based on primary care EHR data in order to determine its validity, concluding that, while generally a valid source, artefacts of variability are observable across EHR databases and different topics, e.g. in that acute diagnoses were not as well recorded as non-acute ones [9].

- Weiskopf and Weng conducted a review of methods and dimensions of electronic health record data quality assessment, identifying different concepts of data quality relevant to secondary analysis [10]; similar concepts were also identified by a work with similar intent by Chan et al. [11], Botsis et al. [12] and Häyrinen et al. in a general, comprehensive review of research literature on EHR system evaluation [1]. Weiskopf et al. reviewed research publications and found, that the data quality trait 'completeness' can have several different definitions depending on the requirements of the secondary analysis effort and accordingly needs to be validated differently [13].

- Herrett et al. reviewed research literature based on secondary analysis of primary care EHR data, identifying several strategies of data validation in order to validate towards data quality concepts [10], as did Weiskopf and Weng [10].

- The potential value of information 'lying dormant' in EHR records has recently been illustrated by Tate et al., who investigated information put in the EHR documentation's free text area and found that diagnoses of ovarian cancer would be documented in free text prior to the date of a diagnostic code [14]. Ford et al. found similar abilities of free text information in relation to first diagnoses of rheumatoid arthritis (RA) [15].

Further related work is cited directly in chapter 7 (*Secondary Analysis of Primary Care EHR Data*) in relation to validation methods and phenomena of variability.

2.3 The Patient Records Enhancement Programme (PREP): The Field Study

This master's thesis builds on data originating from the Patient Records Enhancement Programme (PREP) project. This section describes the background of this project; it complements further detail given in section 3.3 (*Analysis of PREP Field Study Data*).

PREP was a multi-disciplinary research project aimed at finding an "*understanding use of codes in primary care settings in order to develop the best methods for secondary analysis*". Field studies were conducted in order to investigate data entry practices and answering questions of who enters the data, in what context, how, where and why [16]. To this end, consultations with simulated and real patients were video-recorded, walkthrough

interviews with the participating General Practitioner (GP) were conducted and the EHR system usage was recorded with screen recording software.

Two aspects of the PREP field study's have been described in publications: Axelrod et al. (2011a) [16] report on the data recording process. Axelrod et al. (2011b) [17] describe the work with simulated patients in the course of the PREP field study. This thesis takes on recordings of 15 GPs of five UK primary care surgeries, using three different EHR systems, with two simulated patients, for qualitative analysis. The following section 2.3.1 describes the EHR systems (further detail follows in chapter 5), section 2.3.2 outlines the experiment design. Section 4.1 describes the persona profiles of the two simulated patients.

2.3.1 EHR Systems Used by Surgeries Subject to Analysis

EMIS LV. Surgeries HS, EH and SE employ *EMIS LV* an EHR system developed by EMIS Health[1]. EMIS (Egton Medical Information Systems, rebranded to "EMIS Health" in 2015) was founded in the 1980s by Dr. Peter Sowerby and Dr. David Stables, two GPs who sought to create a tool for their practice [18]. Once their flagship tool, "EMIS Health" currently promotes the transition to EMIS LV's successor EMIS Web, reporting in 2013 that a majority of UK GP's have switched to it, either by transitioning from EMIS LV and switching from other EHR systems [19].

SystmOne. Surgery OS employs *SystmOne*, an EHR system developed by TPP[2]. SystmOne features central hosting of its data. Once a unique feature of SystmOne, EMIS LV, by its transition to the equally centrally hosted EMIS Web, moved into a more directly competing role to SystmOne.

Vision. Surgery MS employs *Vision*, an EHR system developed by INPS[3].

EMIS Web, SystmOne and VISION are three major EHR systems utilized in UK primary care, who together cater to the predominant part of GPs in primary care in the United Kingdom. EMIS Web and SystmOne cover around 85% of UK population [20]. They represent three of four *Principal Clinical Systems*[4] approved by the GP Systems of Choice, a national framework installed by UK's National Health Service (NHS). GPSoC formulates policies in service standards and interoperability and allocates NHS funding to the development and deployment of adhering EHR systems[5].

GPSoC-accredited EHR systems are incentivized to improve interoperability and develop interfaces of electronic health record exchange. To this end, several strategies of scaffolding interoperability and aggregation on several levels were implemented, e.g., on a interoperability level, the Medical Interoperability Gateway (MIG), implementing the NHS Interoperability Toolkit (ITK) and HL7 interoperability standards, or, on a data

[1] https://www.emishealth.com
[2] http://www.tpp-uk.com
[3] http://www.inps4.co.uk/vision
[4] http://systems.hscic.gov.uk/gpsoc/interface
[5] http://systems.hscic.gov.uk/gpsoc/about

level, the Summary Care Records (SCR) as a means of central health record keeping. Apart from centrally incentivized interoperability standards, developer companies EMIS Health and TPP also moved to more direct collaboration in 2015, announcing a trial for integration of their databases towards inter-system visibility and shareability [20].

2.3.2 Experiment Design

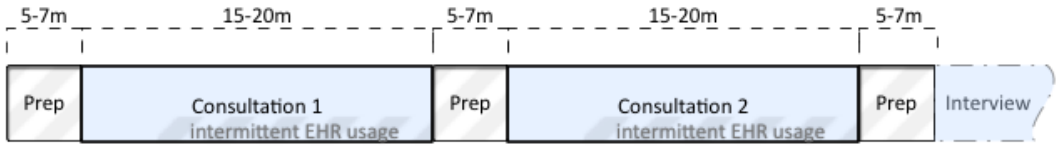


Figure 2.1: Study situation: Consultations were preceded and followed by prep time the GPs used to look at the simulated patient's records and document the consultation.

The consultations with the simulated patients took place consecutively. GPs would take some time to prepare in advance of a consultation to look through the Electronic Health Record (EHR) content of the simulated patient. When ready, they would then call for the patient to enter and hold the consultation. After a consultation, the GPs would take some time to document it in the EHR system. Depending on how much EHR usage would take place during the consultation, this would take between 3 and 10 minutes (see figure 2.1).

This gave the walkthrough interview, which, in all cases, took place after the consultations (with few cases of interviews being held after each consultation) a character of *experience sampling*. Questions in regard of usage practice of the EHR system thus were posed and answered in front of the background of the instantiation of two cases of practice. This, on one hand, is the core advantage and rationale behind experience sampling as a research method, but should be kept in mind as possible limitation as well: The consultations are a lively anchor for otherwise ambivalent and possibly hard-to-answer questions like *"Do you consider any codes or free-text that you then ended up not writing down?"*, but also are an associative frame-of-reference. What is true in terms of GP practice for the two scenarios of a joint pain and sore throat symptomatic might not be the case for other, maybe more severe, maybe more chronic scenarios or any other form of GP work in the broad spectrum of health care that is addressed by primary care. Thus, utterances of the GPs to their work practice should be understood in front of the background of these two scenarios.

2.3.3 Two Simulated General Practice Patient Scenarios

The work with simulated patients has a long-standing tradition in the training of GPs. They provide a "*high degree of realism*" and are used for the assessment of medical and communication skills [21]. This section gives an overview of the profiles of the two

simulated patients as designed by the PREP investigators. Both scenarios were iteratively developed and refined with doctors and with actors [17, p. 76]. They were laid out to resemble common symptom constellations and be "*as contrastive as possible*" [16, p. 565]:

- 1) *Jane*: An older middle-aged woman presenting with a persistent joint pain symptomatic and back-story designed to resemble "red-flag" symptoms and coherently point towards a possible case of rheumatoid arthritis.

- 2) *Scott*: A younger middle-aged man presenting with symptomatic and back-story of a frequently recurring painful sore-throat problem and otherwise minimal context.

Medical Symptoms	Rheumatoid Arthritis	Sore Throat
Gender	Female	Male
Medical history	more	less
"Red flag" symptoms	present	not present
National clinical guidelines	apply	do not apply
Condition type	chronic	acute
Follow-up	likely to be required	may not be required

Table 2.1: "Persona variables" of the two scenarios. [17, Table 1]

Table 2.1 (taken from [17, p. 76]) compares attributes of the two scenarios. (Note that the authors refer to the joint pain scenario as "Rheumatoid Arthritis". In this thesis it is named "Joint Pain", as 1) it is deemed more consistent under a naming scheme by symptoms and 2) it is more grounded in the actual realization of the scenario in the consultations.) The Joint Pain (JP) scenario was meant to convey specifically the onset of rheumatoid arthritis, a chronic auto-immune disease with further follow-up likely to be required. In contrast, the Sore Throat (ST) scenario was laid out to be a comparably unspecific case.

Both simulated patients were equipped with a print-out of an image of (in the ST case) a sore throat and (in the JP case) a hand with swollen, reddened fingers, that they were briefed to show when prompted for an examination (fig. 2.2).

J. Cassell, an academic researcher in the fields of epidemiology and health services research, one of the investigators involved in the creation of the PREP field study who was also interviewed as an expert in the course of this master's thesis, described the intent of these studies as getting a perspective on variability: *"People have built databases to collect pre-specified information and so with EHR becoming such an important information source, we need to know how people are coding in real life, or you'll miss a proportion of people, so we wanted explore diversity of coding, but we also wanted to know, what information doesn't get coded and gets written in free text and written form."*

Figure 2.2: ST scenario persona named Scott Marshall describes his symptoms (left), JP scenario persona named Jane Hallam has 'her' hand examined (right).

2.4 Structure in EHR Systems

In their report commissioned by the National Health Service (NHS) on structure in EHR systems, Morrison et al. [22] investigate the need and utility of structure in EHR systems both for purposes of primary and secondary care. They conclude, among other things, that

> there is a strategic interest in increasing the proportion of the EHR that is captured in structured and coded format" [22, p. 5].

Structure can be imposed on an EHR system in the form of structured entry (e.g., fields for 'Diagnosis', 'Symptoms', 'Examination', ..) and furthermore by the usage of coding schemes, providing the GPs with a vocabulary of terms denoting the information to be put in an EHR system. In this thesis, three coding schemes come up:

Read Read codes are structured hierarchically and aimed at reflecting medical thinking [16]. The Read system widely used in the primary care branch of the United Kingdom. It is the coding scheme used with the three EHR systems subject to analysis in this master's thesis.

International Classification for Primary Care (ICPC) .. a coding vocabulary accredited by the World Health Organization (WHO) for use in primary care. It is the prevalent coding scheme for primary care in Europe [23].

International Classification of Diseases (ICD) The ICD system is a general terminology database of disease. It generally focuses on the diagnostic of disease and is not made out to capture detail as it arises in primary care. E.g., while there are codes for alcoholism in ICD-10, it does not hold the ability to record the amount of alcohol consumed [23], unlike with Read or ICPC.

Further detail can be found e.g. in the work of S. de Lusignan [23], who gives a comparison and historic overview of International Classification for Primary Care (ICPC) and the Read system in light of ICD-10.

CHAPTER 3

Methodology

This chapter constitutes a description of the methodology used in conducting the inquiry central to this master's thesis. Section 3.1(*Paradigmatic Stance*) positions the scientific work in relation to paradigmatic fundaments of qualitative inquiry. Section 3.2 (*Methodological Considerations on Video Analysis*) characterizes the methodology of video analysis as used in this thesis. Section 3.3 (*Analysis of PREP Field Study Data*) describes the analytic process taken to the PREP field study's data in this master's thesis. Section 3.4 (*Literature Review*) describes the literature review done, and section 3.5 (*Expert Interviews*) describes the approach to expert interviews conducted in the course of this work.

3.1 Paradigmatic Stance

The methodology of this master's thesis is set up in line with a constructivist paradigm. Constructivism holds an ontological view that is in its core interpretivist; knowledge is constructed in a process of exchange between the subject of research and the researcher. It is local and specific to the context of these actors. New knowledge is constructed in the form of "*more informed and sophisticated reconstructions [and] vicarious experience*" (Lincoln and Guba, 1994 [24]).

In this sense and also following related work described in chapter 2 ([6], [1], [3]), phenomena of variability and practice in the usage of EHR systems represent 'lived' experience and results of sophisticated constructive processes towards a work practice that allows GPs to function in their roles as physicians responsible to their patients, co-workers related to a surgery and their interests, and not least human beings who bring 'human factors', but also their individuality to the table.

The knowledge this master's thesis seeks to construct is one of practice, a deep understanding of how GPs make sense of the electronic tools provided to them, and what

implications this understanding might hold for the design of such systems. Because of the fact, that the ontology of constructivism does not hold an idea of an objective truth, validation of results lies beyond the reach of a possible direct verification and needs a more nuanced approach also depending on the utilized methodology. This thesis takes a qualitative approach to both the PREP field study's data as well as the expert interviews; the method of analysis is *Thematic Analysis*, an approach described by Braun and Clarke [25], that is related to Grounded Theory, to the end of identifying patterns of variability and influencing factors of practice in the work with EHR systems that might be relevant when the EHR data is to be used for secondary analysis (described in more detail in section 3.3).
In conducting Grounded Theory (in this work, substituted by the closely related Thematic Analysis, because of its focus on themes rather than forming a theory around the observed phenomena), Charmaz argues for a constructivist stance:

> "[A] constructivist view assumes an obdurate, yet ever-changing world but recognizes diverse local worlds and multiple realities, and addresses how people's actions affect their local and larger worlds. Thus, those who take a constructivist approach aim to show the complexities of particular worlds, views, and actions". Charmaz, quoted after Krüger and Meyer [26, p. 13]

Validation of constructivist knowledge can follow different quality criteria. Lincoln and Guba name *trustworthiness* and *authenticity* [24]. Charmaz suggested the following criteria for evaluating Grounded Theory studies: *Credibility*, *originality*, *resonance* and *usefulness* [27]. In this sense, this master's thesis is set out to achieve credibility in both its groundedness in the practice of GP work, the invitation of experts to comment on the results and a thorough literature review to situate the results in a meaningful manner in what relevance the state-of-the-art might attribute to it, in order to yield knowledge useful to the research community.

3.2 Methodological Considerations on Video Analysis

Visual data is a versatile means of scientific practice amenable to numerous methodologies [28, paragraph 3] and referred to in diverse terminology, e.g. "video-taping", "video-technology", "video-analysis" or "video-observation" [29, p. 29]. This chapter is based on information from the works of Knoblauch [28] [30], Jewitt [31], Goldman and Ricki [32], Garcez, Duarte and Eisenberg [33], Lomax and Casey [34] and Mruck [35].

The concept of video analysis applied in this thesis is a qualitative, ethnographical one; it draws from Hubert Knoblauch's concept of Videography: A term suggested by him as a portmanteau of video analysis and ethnography, *Videography* is a methodology of analyzing video-recorded human behavior in social settings, an "ethnography utilizing video" [30, quoted after [29, p. 29]]. Videography is a means of *focused ethnography* [36] directed at specific aspects of interest, like the way GPs interact with computer

technology during, before and after medical consultations. This relates to the ability of video recordings to collect a large amount of data in comparably short time.

Video analysis, as any research method, has its reach, strengths and weaknesses. It allows for the observation of human interaction in its verbal and nonverbal layers, e.g. setting it aside from audio-recordings. It collects a dense stream of data and holds the ability to provide a rich, highly resoluted picture of a lively, dynamic reality, of *"fleeting and non-repeatable events which are very likely to escape direct observation"* [37, p. 423], not least because data collected via video-recording can be effectively reproduced and analyzed in detail, e.g. by time lapse, slow motion and repeated watching. This also makes video analysis a suitable tool to analyze people in their interaction in social situations, since such activity is influenced by the simultaneity of potentially several layers, e.g. social interaction with others as well as individual acting [38, quoted from [29, p. 38]], a density that can be made more easily accessible with video-recording.

Video data furthermore is a suitable type of data to be approached by "*another set of eyes*", as e.g. noted by Knoblauch, saying that video-analysis can be successfully undertaken also by persons not involved with the data collection, enabling a detached, impartial perspective [39, p. 130]. Jewitt attributes to video data the notable abilities to be shared, re-opened for more analytical passes and generally "*to support an exploratory research design or data-discovery phase*" because it retains its informational density longer than other data collection methods and is focused down only in later stages of analysis [31, p. 7]. Knoblauch argues that

> [these technologies] do change the character of the data produced: one is no longer subject to the "incurable subjectivity" of field reports but disposes, instead, of a substantial degree of intersubjectivity. This does not mean that recorded data are more objective; it does, however, allow for outsiders to access the data which are less dependent on subjective perspectives than are field-notes. [36, paragraph 23]

Reality and Data Collection for Video Analysis

Notably, video recording leaves out information, e.g. compared to participating observation in that the personal experience, smell and touch experienced by the researcher in situ can hold or contribute to insights as well. The perspective of the video-camera is a partial one in that its field of vision is deliberately chosen by the camera's operator. Efforts of analysis of video-data need to take into account that the high density and perceived realness of video-data may be deceiving. Video data, not unlike numbers and statistics, arguably has a questionable persuasion of exactness, of objective reality, attached to it.

An important caveat in interpreting video data is considering possible influences of the *Hawthorne Effect*[1], signifying a change in behavior in individuals resulting from their

[1]Exemplary definition e.g. available from Encyclopaedia Britannica, http://www.britannica.com/topic/Hawthorne-research, accessed 15-November-2015

Figure 3.1: Perspectives can be deceiving. (Source unknown.)

awareness of being observed. The presence of equipment demonstrates the presence of an observer [36, paragraph 22], and a recording situation can inject artificialness into the supposedly untarnished reality, as e.g. suggested by Lomax and Casey: Individuals who appear at ease and immersed in their recorded activity might actually be interpreted as actively not paying attention as they might understand their role as "*one in which doing ignoring was appropriate*". [34, § 2.2].

The influence of feeling observed in the course of video recording on the portrayed reality is approached in different ways:
If considered a distortional effect, the validity of results can be sought to be improved by "*technical and non-technical strategies*" [31, p. 9], like more and more clandestine camera equipment up to the point of research being "done covertly, which immediately raises traditional ethical questions whatever the nature of the research" [34, paragraph 3.1] or the employment of "*respondent validation techniques*", e.g. post video-taping interviews to the end of inquiring whether and how the individual felt to have behaved differently [34, paragraph 3.3]. A flawed strategy, according to Jewitt, because such interviews presume that "participants have privileged insight into the social interactions of which they have been a part" [31, p. 9]. Another stance on this issue often taken in ethnographic research [31, p. 10] is a *reflexive* one which understands behavior not only as an act by itself, but also as related to interpretational perception on the side of the recorded individual, making behavioral changes in face of a video camera a point of investigation and possible source of insight in itself. Lomax and Casey as an example describe their video-study of midwifery practice [34, paragraph 4.3], in which the participants' reaction to being filmed and their interaction with the researchers on what was relevant to their practice or should be left out was examined as evidence of their self-understanding in their profession and enactments of boundaries in their work.

3.2.1 Data Preparation and Analysis

Video data can be handled in various ways prior to or during its analysis: Time-lapse and slow motion etc., as mentioned in step 1, can provide a detailed view on recorded events. Transcriptions and still frames can focus on certain aspects of a recorded scene, e.g. spoken words, mimics and/or gesture. Video data furthermore can be broken down into parts, which in turn can be recombined alongside processes of analysis [40].
Video analysis is supported by computer technology and software tailored specifically to the needs of scientific inquiry on video recordings. Garcez, Duarte and Eisenberg report on their usage of *Atlas.ti*, a computer software product aiming at supporting qualitative analysis of video data, in an effort of data analysis in qualitative research [33]:

> "The process of encoding data and the construction of categories in content analysis of videotaped material using computer programs is similar to that used in the analysis without a computer: the researcher sees all the scenes previously cut and encoded (units of analysis), and associates them with categories and / or key concepts and / or thematic areas (depending on his/her theoretical and methodological affiliations) defined by him or her, taking into account the research guiding question(s).
> In the next phase, with these sets one can build maps and / or detailed analytical frameworks, from which one can make the global description / interpretation of the empirical materials in dialogue with the theoretical framework.
> The main difference in the use of computer programs is that citations (cut scenes / units of analysis) and their encodings can share the same operating environment at the same time. This enables the researcher to view classified material as often as s/he deems necessary and reassess the adequacy of the categorizations that s/he produced. Thus, s/he can not only create not new categories and / or delete, replace, merge, modify, link existing ones, but also relocate citations. From our point of view, the use of a suitable program can bring greater agility, flexibility and transparency to the process of analysis." [33, p. 258].

The analysis of video data is a detail-oriented activity that works hand-in-hand with a general analytical approach framing and driving it. A common combination also recommended by Knoblauch [39, p. 136] is to complement 'small-scale' video analysis as a means of interpretation with *Grounded Theory* as a method of selection and combination of data and inductive theory formation.

The choice of analytical approach in this thesis is made under consideration of the methodology of related work on ethnographical field studies with GPs and is described in more detail in the following chapter 3.3.3. For the purposes of this chapter on general, methodological considerations on video analysis, it should suffice to say that many ways to approach the 'small-scale' analysis of video recordings exist, differing depending on the research paradigm at hand and respective research question(s) as well as the video-recorded phenomenon of interest, without going into further detail.

3.3 Analysis of PREP Field Study Data

This chapter describes the preparation and analysis undertaken with the Patient Records Enhancement Programme (PREP) study's data in the course of this thesis. For information on the research design and data collection of the PREP field study, the origin of the data analyzed here, see chapter 2.3 which details on its background.

Ethical Aspects

Handling participant or person-related data requires an attentive stance to ethical aspects. The PREP project's field study was conducted following approval by UK's National Health Service (NHS) ethics board and the ethics committee of the University of Sussex [16]. The GPs and patient actors who participated in the PREP project's field study gave consent to procedures of data-recording and usage as well as usage of the information disclosed by them.

In the course of this thesis work, measures were taken to ensure the substance of this consent would be upheld without compromise, as the handling of such sensitive consultation and personal data requires. The author of this thesis signed a non-disclosure agreement in order to be allowed to review the data. Subsequently, depictions of GPs and patient actors, when used as examples in the expert interviews (see section 3.5) or in this thesis, were anonymized. Personal details like log-in names in EHR systems were removed from screenshots.

3.3.1 Overview of GPs and Surgeries

The investigated data encompasses recorded consultations of 15 GPs of 5 surgeries, which are referred to as HS, OS, EH, MS and SE in this master's thesis. All surgeries are located in the United Kingdom; surgeries OS, MS and EH in a rural location, surgery SE in a semi-urban location and surgery SE in an urban location.

Surgeries HS, EH and SE use EMIS LV, surgery OS uses SystmOne, and surgery MS uses Vision as EHR system. Figure 3.2 gives an overview of this data, including age range and gender of the involved GPs. All GPs conducted consultations with both simulated patients (described in section 4.1), yielding 30 recorded consultations as basis for analysis. For further details to the utilized recordings and documents, see appendix section A.1.

3.3.2 Data Preparation

The PREP field study data analyzed in this thesis comprises (1) video-recordings of 30 GP consultations, conducted by 15 GPs working in five surgeries, using three different EHR systems, (2) screen-capture videos of the EHR system usage during these consultations, (3) video-recordings of walkthrough interviews conducted with the 15 GPs by the field study's investigators and (4) detailing documents and observational field notes on the five GP surgeries as the GPs' places of work. The video material all in all consists of 20 hours of consultations, walkthroughs and screen-capture videos.

Surgery	EHR System	GPs (gender, age range)			
Surgery HS	*EMIS LV*	**GP1** female, 30s	**GP2** male, 60s	**GP3** female, 20s	
Surgery OS	*SystmOne*	**GP4** male, 40s	**GP5** female, 30s	**GP6** female, 40s	**GP7** female, 30s
Surgery EH	*EMIS LV*	**GP8** female, 40s	**GP9** male, 40s	**GP10** male, 50s	
Surgery MS	*Vision*	**GP11** female, 30s	**GP12** female, 30s	**GP13** female, 20s	
Surgery SE	*EMIS LV*	**GP14** female, 50s	**GP15** female, 40s		

Figure 3.2: Overview of GPs and surgeries: GPs with gender and age range, surgeries with employed EHR system.

3.3.2.1 Video Recordings of Consultations, Screen Captures and Walkthrough Interviews

Consultations This thesis examined recordings of 15 GPs during their consultations with the two simulated patients, resulting in 30 consultation recordings. The recordings average at 25 minutes in length for one consultation.

Screen Recordings To 10 of 15 consultation recordings, screen recordings were available. These recordings were made using screen capture software and depict the usage of the EHR system during the in real time, showing the GP's usage of codes and free text in the course of her/his consultation.

Walkthrough Interviews Following the consultations with the simulated patients, the field study's investigators conducted walkthrough interviews with the attending GPs. Recordings of these interviews with the GPs of all 30 examined consultations were examined in this thesis. They average at 20 minutes in length, were included in the video analysis and examined with a separate coding set.

The GPs would hold both consultations and then be interviewed, mostly in the order depicted in figure 3.3, resulting in video files of mostly one continuous recording of both consultations and the walkthrough interview.

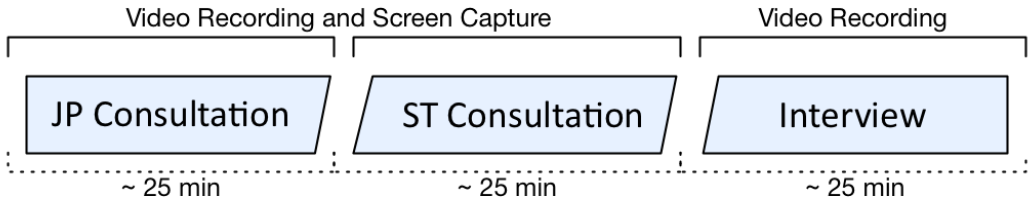


Figure 3.3: One recording session with a GP consists of the Joint Pain and Sore Throat (**JP**, **ST**) scenario consultations and a walkthrough interview afterwards.

3.3.2.2 Detailing Documents

In addition to video data, documents detailing on several aspects were available:
Field notes: For all surgeries, observational field notes on the study sites.
Simulated Patients Profiles: A document describing the persona profiles of the simulated patients.

3.3.2.3 Atlas.ti

All data was imported in the *Atlas.ti* analysis software. There, the video files were segmented into multi-layered cut scenes (referred to as *quotations* in *Atlas.ti*) and used in an analysis process described in the next chapter 3.3.3 (Data Analysis).

3.3.3 Data Analysis

This section describes fields of interest in the PREP field study data, gives an overview of the analytical process and the research questions and then details on its methodological stages.

3.3.3.1 Fields of Interest

The constructive process of the consultation: To understand the content of the consultations, it is relevant to examine the constructive process that unfolds in the course of a consultation. Roberts and Sarangi approach this constructive process by discourse analysis and suggest a theme-oriented approach to the end of identifying relevant *focal themes* (relevant to the profession, for example topics of conversation) as well as *analytical themes* (relevant to the process, for example contextualization, inference or framing) [7]. To examine this, Swinglehurst, Greenhalgh and Roberts conduct conversation analysis in a social constructivist perspective

> that is to say language does not just reflect or express intentions or decisions (the representational role of language), but makes them (the constitutive role of language) - talk is work. [1, p. 3]

Organizational routines GP consultations take place in front of the background of the medical profession and its organizational routines. Feldman and Pentland describe organizational routines as a conceptualization of work on the side of the professional that is in a constant state of change and appropriation. It exists as an abstract idea (*ostensive*) as well as its practical enactment (*performative*) and shapes the *artefacts* that codify and mediate this work [6]. This represents tacit knowledge of interest in this work. Swinglehurst, Greenhalgh, Myall and Russell use the concept of organizational routines as 'unit of analysis' in an ethnographical case study of general practice [17].

EHR systems can play an important role both in the constructive process of the consultation and the evolution of an organizational routine. Their effect is observed on its own and in relation to both described fields of interest and is meant in line with what Swinglehurst et al. refer to as sociotechnical approach:

> [A] focus on the dynamic, contingent interaction between humans and technologies rather than assuming technology is itself 'causal' of specific effects. [1, p. 2]

3.3.3.2 Overview of Analytical Process

The analysis of the videographical data in this thesis was conducted in an inductive process of Thematic Analysis. In line with the steps of video analysis described by Denzin [41] (see chapter 3.2.1) and the methodology of Thematic Analysis [25], this comprised a phase of immersion in/familiarization with the video material followed by an iterative viewing cycle of more and more refined analysis and coding, conceptualizing, categorizing and reasoning towards theme generation in relation to the research questions.

Embedded in this "large scale" approach were different small-scale takes on the two different types of video data:

Consultations The video-recordings of the consultations (and the respective screen captures) were analyzed on the basis of the conversation-analytical view on consultations described by Swinglehurst et al. [1]. In the course of this analysis, the recordings were subdivided into quotations, which were subsequently annotated and mapped following the process of Thematic Analysis. In some cases, transcripts were created to better compare interactions as well as to serve as quotations in this thesis.

The consultations furthermore provided a view on the performative side of the organizational routine enacted by the GPs as well as the artefacts produced in the course of the consultation in the form of EHR system entries (as exposed by the screen capture videos).

In the quest to find and describe forms of variability in GP work, the GPs' interaction with the EHR system again was of interest as a tool assisting the constructive

process during the consultation, as a medium of artefact generation in Computer Supported Cooperative Work (CSCW) and by itself as an influencing factor to GP work in general.

Walkthrough Interviews The walkthrough interviews were analyzed as semi-structured interviews in an explorative manner. As described in section 2.3, the interviews were conducted by PREP field study's investigators around questions of the specific EHR usage during the two scenario consultations as well as usage practice of EHR coding and free-text in general.

In this sense, the walkthrough interviews were of interest for yielding a view on the ostensive aspect of the organizational routine enacted by the GPs, for explanatory words on the EHR codes and free-text idiosyncrasies and EHR usage practice in general.

The analytical process is familiar to the one used by Swinglehurst et al., who described an iterative four-stage process of a) familiarization, b) multimodal transcription, c) micro-analysis and d) identification of themes. [1, p. 3, fig. 1]. The analytical process applied in this thesis differs in that not transcriptions of the consultations and interviews were subjected to analysis, but video scenes themselves. This follows the argument of Garcez et al., who argue that video should not be transcribed as this is a lossy conversion in terms of richness and potential [33, p. 259]. This approach is furthermore made possible by *Atlas.ti*, the video analysis software employed in this analysis. Its capabilities of efficiently subdividing video clips into multiple layers of quotations as well as structuring and reorganizing them by codes, code groups and a combinatorical visual analysis in graph networks allow for a qualitative reasoning process without the need of reduction into a written transcript.
Figure 3.4 gives an overview of this process. Swinglehurst et al.'s step of multimodal transcription is replaced by a step of structuring.

Research Questions

To guide the analysis of the PREP field study's data, research questions were put in place that aim at identifying possible forms of variability in consultations and EHR documentation across the three different brands of EHR system employed (Vision, SystmOne and EMIS LV) as well as across the surgeries:

- RQ 1: *How are the two simulated scenarios dealt with in the consultations and how is the information originating from the consultations documented in the EHR systems at hand?*
- RQ 2: *How does this relate to the understanding GPs have of the building blocks of EHR documentation - coded entries and free text?*

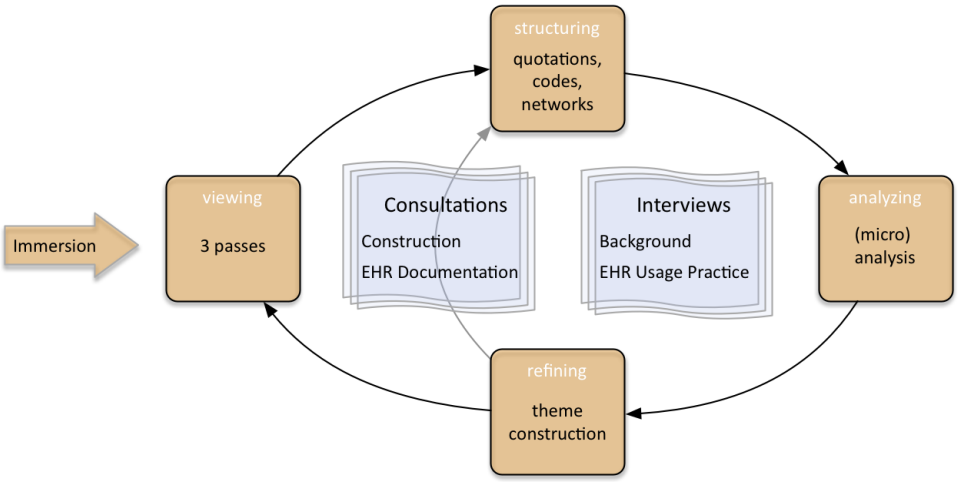


Figure 3.4: Overview of the analysis process applied to the data set. Three passes of viewing the consultations and interviews yielded coding schemes in different categories towards themes in line with a grounded theory approach.

This to the end of describing usage practice in general as well as the effect possible forms of variability have on the resulting EHR data in order to better understand how and why EHR documentation comes to be.

3.3.3.3 Video: Analysis Process Towards Themes

In the course of the multiple passes of viewing, structuring, analyzing and theme construction (figure 3.4) of the PREP study's field data, the analysis process can retrospectively be subdivided in three phases; (1) initial immersion and skimming of the recordings, (2) three passes of full viewing of all recordings (see fig. 3.4 to identify, substantiate and find illustrating examples for themes, followed by (3) a phase of fine-grained analysis en passant the report generation.

Viewing Pass 1: Quotations

After an initial subdivision of the video-files into consultations and walkthrough interviews done during the immersion phase, the analysis during the first full viewing pass was focused mainly on subdividing the 30 GP consultations and 15 walkthrough interviews into scenes of interest. These quotations average at 2-3 minutes in length and were created both as a first pass of open coding and in relation to the fields of interest laid out early on in the analysis process. While, in part, quotations were created in later stages of analysis as well, most of the work in the first viewing pass was done on quotation level and vice versa.

The **consultations** were subdivided into 581 quotations on multiple levels, roughly equally divided among all 30 consultations. The initial quotation-making as a stage of

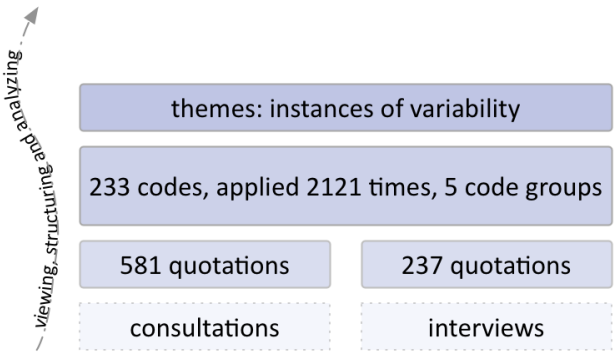


Figure 3.5: Progression of the analysis process towards theme refinement.

open coding identified scenes of interest in an explorative and candid approach. This first level of structure yielded mostly *focal themes* that occurred during the consultation, e.g. when a GP conducts an examination, asks for an account of the present symptoms or related personal history of the simulated patients or issues prescriptions.

The **walkthrough interviews** were subdivided into 237 quotations following the thematic flow of the interview. Initially, mostly interesting quotes in relation to EHR usage practice were marked by quotation - a process that was later on replaced by quotation-making along the interview questions which reoccur in varying stability.

Viewing Pass 2: Coding and categorizing

An interim analysis of the structure yielded by the quotation set implied certain codes and code groups that constituted a starting point for the coding effort in viewing pass 2. Four code groups developed as basic categories:

GP and Consultation The focal themes occurring in the consultations that were found in viewing pass 1, were abstracted into concepts reoccurring throughout all consultations, like the establishment of a symptom set, of past history, of a concept of "what is going on" and the derivate of a treatment plan. These, along with behavioral themes, were subsumed under this basic category.

EHR Instances of EHR usage during the consultations as well as descriptions of ostensive aspects around EHR usage were subsumed under this basic category.

Simulated Patients Interesting aspects related to the simulated patients, e.g. around how they unfolded their scenario during the consultation or aspects that could be seen as limitations e.g. in believability originating from their "simulatedness" were subsumed under this basic category.

Walkthrough Interviews Interesting snippets as well as reoccurring conversational themes around GP practice were coded under this basic category.

During viewing pass 2, the semantic separation of consultation and walkthrough interviews gradually dissolved towards an overarching coding scheme. While consultation and walkthrough interviews still have a designated basic category, many codes were used across these categories. Both the basic categories of *simulated patients* and *EHR* existed beyond the consultation/walkthrough separation. In the course of viewing pass 2, 233 codes were defined related to these four categories and further subcategories. They were used 2121 times.

Viewing Pass 3: Report generation

In the course of the generation of the reports on the video analysis of the consultation recordings (see chapter 4) as well as on the themes that emerged from the analysis of the walkthrough interviews (see chapter 6), another full viewing pass was conducted.

3.3.3.4 EHR Documentation: Comparative Content Analysis

To identify themes of variability in the artefacts produced in the course of the consultations, a content analysis of the EHR documentation was conducted. This analysis was guided by RQ1 (see section 3.3.3.2 on research questions). EHR entries were thematically tokenized (e.g. a description of *Jane*'s Family History would be denoted by <Jane's Family History>) and, on that basis, EHR usage was compared across GPs, surgeries and EHR systems along the usage of Read codes and free text categorization.

This approach was combined with an examination of how the three EHR systems, EMIS LV, SystmOne and Vision, would be designed in three key aspects:

- How are patient overview visualizations structured?
- What means of structuring an EHR entry are available to the GP?
- How is the utilization of Read codes integrated in the creation of an EHR entry?

Chapter 5 constitutes the report on the results of this approach.

3.4 Literature Review

To provide background on the topics relevant to this master's thesis, a review of the research literature has been conducted. To this end, publications in fields of computer science as well as medical and statistical science were reviewed along the following questions:

EHR-Design What current challenges in EHR design exist - how are EHR systems designed and validated?

Secondary Analysis What usage scenarios exist around the secondary analysis of primary care EHR data? What data-related procedures constitute a state-of-the-art in assessing the quality of data used for secondary analysis?

3.5 Expert Interviews

In order to relate results of the PREP field study's analysis to aspects of secondary analysis, semi-structured interviews were conducted with six experts from related fields. They were selected on basis of their involvement with secondary analysis (experts 1-3), their involvement with the PREP project (expert 4, also experienced with secondary analysis), their expertise with EHR systems (expert 5) and Natural Language Processing (NLP) (expert 6). They were approached via E-Mail invitation.

- 3 epidemiologists with different fields of interest: Primary care, Clinical cancer research, Virus epidemiology/influenza surveillance. They are academic researchers of the Medical University of Vienna, Austria.
- 1 epidemiologist and health service researcher: Focus on primary care and aspects of evaluation of health services. Academic researcher based in Sussex, United Kingdom.
- 1 expert on the design of Electronic Health Record (EHR) systems. Academic researcher of the Medical University of Vienna, Austria.
- 1 expert on Natural Language Processing (NLP) in EHR systems. Academic researcher of the Medical University of Vienna, Austria.

Expert	Field of Expertise
E1	Epidemiology - Secondary Analysis of Primary Care Data
E2	Epidemiology - Mortality, Co-Morbidity of Cancer
E3	Epidemiology - Viruses - Influenza Surveillance in Austria
E4	Epidemiology, Health Services Research - Secondary Analysis of Primary Care Data, PREP Project Investigator
E5	EHR System Research - Semantic Interoperability, ELGA
E6	Natural Language Processing - Information Retrieval

Figure 3.6: List of experts interviewed in the course of this work and their reference (E1-E6) in the thesis.

Interviews with experts E1, E2, E4 and E6 were audio-recorded, the interview with expert E3 was documented by the interviewer taking notes during the interview. Prior to the interview, the participants signed a consent form, which gave them information on the process and how the resulting data would be treated in the course of this thesis work (see appendix B.3 for further detail).

3.5.1 Interview Design

Semi-structured interviews aim at approaching topics of interest not by explicit questions as in questionnaires, but utilize open questions in order to, in part, hand over control of the conversation to the interviewee; this to the end of approaching a topic exploratively and gathering information alongside what U. Flick refers to as the *subjective theory* of the interviewee:

> Interviewees have a complex stock of knowledge about the topic under study. This knowledge includes assumptions that are explicit and immediate and which interviewees [..;] [t]hese are complemented by implicit assumptions. [42, p. 156]

In describing the expert interview as a special form of semi-structured interview, Flick highlights its value as a supporting method:

> The systematizing expert interview can be used to collect context information complementing insights coming from applying other methods (e.g., interviews with patients). [42, p. 166]

Expert interviews differ from semi-structured interviews, among other ways, in a time-pressure on the side of the interviewee because of their function in their field. With this in mind, and further mindful of the Flick's remark, that expert interviews do not only serve the purpose of gathering abstract specialist knowledge, but to access practical knowledge [42], fields of interest for the interviews with the six experts were conceived:

- Establishing the background - research interest in epidemiology / EHR research / NLP research
- Reflecting on the PREP field study analysis results - would the experts expect this variability and how do the judge the relevance of it in respect of their practical experiences with secondary analysis of EHR data?

Further fields of interest were explored depending on the expert's background:

Epidemiologists *Which data do they use for their research? Where do they get data for their research? What are validity problems of such data? What is necessary for them to feel confident about their data?*

EHR Expert *Do different EHR systems yield different EHR documentation? Would interface design guidelines improve EHR systems in terms of reducing variability? What are challenges for NLP in EHR data?*

NLP Expert *What are challenges for NLP in EHR data? How relevant is data structuring for its quality in respect of information extraction via NLP?*

The interview outline is attached in appendix section B.

3.5.2 Analysis

The recordings of the expert interviews (in case of E3, the notes taken during the interview) were subsequently analyzed by means of thematic analysis (see previous section); this analysis was directed at identifying prevailing views on data coming from primary care EHR systems as well as possible phenomena of variability influencing it.

CHAPTER 4

Jane and Scott in Consultation

This chapter constitutes a report on the video analysis of the consultations with *Scott* and *Jane*. Section 4.1 (*Overview of Scenarios*) gives an overview of their scenarios and how they played out during the consultations. Section 4.2 (*Lifelikeness and Consistency of Simulated Patients*) describes reflections on strengths and weaknesses in the performance of the simulated patients. Section 4.3 (*Basic Focal Themes in a Consultation*) recounts the course of the consultations along common basic focal themes. Section 4.5 (*Summary and Conclusion*) concludes this chapter and formulates key take-aways.

In the interest of a lean reading flow, remarks related to the Joint Pain scenario will be referred to as *Jane*'s scenario and the Sore Throat scenario will be referred to as *Scott*'s scenario, based on the enacted patients' fictitious names.

4.1 Overview of Scenarios

Following the description of the general methodology of simulated patients in chapter 2.3 and based on the recorded consultations, this section gives a more detailed description of the two scenarios of *Jane* / the Joint Pain scenario and *Scott* / the Sore Throat scenario and how they were enacted.

Medical Symptoms	Rheumatoid Arthritis	Sore Throat
Gender	Female	Male
Medical history	more	less
"Red flag" symptoms	present	not present
National clinical guidelines	apply	do not apply
Condition type	chronic	acute
Follow-up	likely to be required	may not be required

Table 4.1: Persona variables of the two scenarios. [17, Table 1]

Table 4.1 outlines the basic idea and intention of the two scenarios. This work uses "Joint Pain" to denote the scenario named "Rheumatoid Arthritis" by the PREP investigators, because this term better encompasses what *Jane* presented with in the consultations and is more situated in the actual realization of the scenarios.

4.1.1 *Jane* / The Joint Pain Scenario

Jane Hallam, as the JP scenario's persona was named, is a woman in her 40s situated in a higher socio-economic band. She leads a healthy lifestyle in that she never smoked and is a very moderate drinker. She is not overweight and exercises moderately. She has no allergies. She works as an accountant and has one child. Her grandmother had some form of arthritis, but she doesn't know more.

Jane presents with a case of persistent joint pain in hands and feet. She had been using a mild pain medication on occasion. *Jane* described her symptoms to the most part consistently as in quote 1. With exceptions (GP11: 3-4 weeks, GP6: 4 weeks), Jane reported this condition to be persisting since about six weeks.

GP: What can I do for you?
J: Well.. I've got this, sort of erm, pains in my joints.. in my feet and my hands
GP: **mhm**
J: that's been going on a few weeks now so ..
GP: - *ok* -
J: - that's why i've come.. to get it checked out.
GP: And how did they start?
J: Well it was about six weeks ago.. it just started as.. sort of just a bit of twinging.. and you know. just generally aching -
GP: **nods**
J: And it has just kind of got worse and worse
GP: *ok* - Any particular time of day?
J: Very bad in the morning -
GP: *Yeah*
J: First thing in the morning, I really struggle with it -
GP: The pain?
J: - Stiffness.. and generally pain -
GP: *Yeah*
J: It's definitely worse in the morning -
GP: And it's mainly hands?
J: - Hands and feet -
GP: Any other joints involved?
J: No, just the hands and feet.
GP: Ok. Otherwise fit and well?

Quote 1: Jane describes what leads her to GP1.

Jane's story, being one designed to resemble the onset of rheumatoid arthritis, also contains more subtle details that might ring a bell on the side of a GP, like the fact that she had experienced a transitory bout of similar symptoms around the birth of her child. This, in line with the miscarriage, which is an event also exhibiting some correlation with rheumatoid arthritis, forms a scenery that would not be explicitly explored by most GPs during the consultation. *Jane* sometimes would still weave it in (GP7, GP8, GP9, GP10), as shown in quote 2, which is also an example for the dynamic unfolding of the scenario depending on the GP's questions.

This information was also made available to the GPs via the EHRs: *Jane* has a medical history that was placed in the EHR systems in advance of the simulated consultation, as illustrated in figure 4.1.

Date	Description	Priority	Clinician
15/10/09	Ca cervix screen normal		----
08/09/09	Ca cervix - screen done		
29/08/07	Hd Ideal weight 62.7	5	RD
	Weight 62.73 kgs BMI: 23 O/E - weight		SSE
01/03/07	Oral contraception NOS First service Claim Expiry Date:		RD
24/08/06	Weight 60.91 kgs BMI: 22.3 O/E - weight		----
	Height 1.651 metres O/E - height		SSE
02/04/06	Hs Nausea intermittent - no stomach pain or diarrhoea/ ? viral	5	
22/05/05	He Cervical smear: negative Normal	5	
12/04/05	He Cervical smear screen Taken	5	SW
11/09/04	Ca cervix screen normal		RD
06/05/03	P/N - tenth day visit at days Clinician: Dr Smallfield Surgery all well		SSE
27/04/03	Hi Normal delivery girl	5	
12/05/02	He Ankle sprain left. Swollen but not broken. Suggest take paracetamol for pin	5	
25/10/01	Hd Miscarriage patient coping support from family. Nothing of concern	5	
17/06/00	Hs Married	5	

Figure 4.1: *Jane*'s medical history placed in the EHR systems prior to the consultations. (GP11. Surgery MS. EHR-System: Vision.)

In all but a few consultations (GP13, GP14), *Jane* would be asked for and mention occasional self-medication with Nurofen, an over-the-counter pain medication, to help her cope with the symptoms. In some others (GP2, GP6, GP9, GP10) she would underline the severity of her situation with mentioning that she had to stop going to aerobics class because of too much pain. *Jane* would furthermore frequently (all but GP12, GP14) ask whether it would make sense to use a wristband as a support during her computer work in her job.

GP: So, just explain to me, sort of, a little more how the happenings have progressed.
J: Well, erm, a few weeks I first started noticing these aches and pains, but I did have a similar thing just after giving birth to my daughter for a couple of weeks.
GP: **nods**
J: So, to begin with, I just thought it's the same thing and it went away that time.
GP: - *right* -
J: So, I felt generally achy and pain.. and then it's got worse, and now it's sort of.. kind of.. all the time.
GP: - **nods** -
J: I've got twinges and .. particularly in my hands and feet.. it definitely.. isn't getting better
GP: - *no* -
J: it's getting worse.
GP: Ok. Have you tried anything to help with these aches and pains?

Quote 2: Jane weaves in details (GP1).

To all GPs, *Jane* mentions her fears in relation to the future course of her condition, as she remembers her grandmother and how she was badly affected by arthritis, "*sort of bent double*".

4.1.2 *Scott* / The Sore Throat Scenario

Scott Marshall, as the ST scenario's persona was named, is a man in his 30s situated in a lower socio-economic band. He is slightly overweight and consumes alcohol regularly to some extent. He is an ex-smoker. He works in a call center and has three young children. *Scott* has a medical history of allergy (hay fever). Allergies and skin problems are also present in his family (father, sister).

Scott presents with a case of acute, painful sore throat that had been recurring three times in the year past. He uses pain relief and various cough sweets.

Initially having been intended to resemble a common situation of a patient presenting with a fairly standard sore throat problem, this scenario ended up being enacted slightly differently: While, intentionally, his core medical problem was supposed to be a sore throat and elements of the story pointing towards a stressful life (three young children, commission-based work) were to help add to a more vivid impression, the actor of *Scott* took these environmental stress factors into the core of the medical problem, changing the character of the scenario. In addition to describing his sore throat problem, he would appear rather tense and worried and - often repeatedly during one consultation - describe his fears that the recurring sore throat problem will cause financial problems for his family (quote 4).

In most of the consultations, *Scott*'s self-medication would come up: He would mention cough-sweets (GP2, GP6, GP7, GP11, GP13), Nurofen (GP4, GP6, GP7, GP14, GP15), Paracetamol (GP3, GP9, GP10, GP13) or that he didn't take any medication (GP5).

In roughly half of consultations, him being an ex-smoker would come up (GP2, GP8, GP9, GP11, GP12, GP13, GP15).

Scott would describe his reason for consulting with the GP consistently as in quote 3 and quote 4.

GP: What can I do for you?
S: Well yeah I just, erm, I've been having this sort of, erm, really, really sore throat, you know, just, erm, really painful.
GP: - *mhm*-
S: I've got this cold as well. But, I mean, it's been going on now for, I mean, I've had it for about three times last year.
GP: - *mhm*-
S: I've had it a few months ago, and it has come back again. It just keeps coming back. It's just so.. painful. You know, all the time. Erm, you know, so I thought I should check it out -
GP: - *okay* -
S: - see what's going on you know, so.. that's the size of it, really, I suppose, yeah, it's just the throat. Yeah.
GP: So, when did it all start this time?

Quote 3: Scott describes what leads him to the GP.

GP: Do you have a theory as to why this is coming on?
S: Erm, yeah, I think I'm under quite a bit of stress, because my work is commission-based, so, it's a bit of a vicious cycle -
GP: - *mhm* -
S: - because if I don't go into work I don't get paid, and that's stressful, so, they kind of pressure me to be there. I've got three young kids, you know, -
GP: - *yes* -
S: - twins of three plus a five-year-old -
GP: - *gosh* -
S: yeah, I suppose stress is not helping.

Quote 4: Scott describes his sorrows.

4.2 Lifelikeness and Consistency of Simulated Patients

As mentioned in chapter 2.3, the method of sending actors trained on a scenario into a consultation was chosen, among other reasons, to draw realism from the human ability to act, ad-lib and dynamically adapt to the GP's questions [17] to the end of making a simulation more life-like. As a corollary, some inconsistencies exist in the scenario exposition from consultation to consultation along such ad-libbing.
Observably, the task of enacting a credible patient in light of the anamnestic questions of and interaction with the GP is one demanding for high concentration in the actors, as they are on constant look-out for opportunities to weave in particles of their story, as in quote 2, or for situations that would take them beyond their script. In such cases, the actors were advised to be vague in their reaction, as this would both be possible to them and also "*often the case in real life*" [17, p. 78], e.g. quote 5.

When asked for feedback on them (GP1, GP3, GP8, GP9, GP10, GP15), the GPs would describe their performance consistently as credible, very good, well prepped and convincing, *"exactly what we do in our GP training"* (GP3). Two GPs described the consultation situation as stressful, e.g. GP4: *"I was incredibly nervous.. you're being*

observed at something you're supposed to be good at" (GP4). The actors would mix-up their story a little from consultation to consultation, but with little significance, e.g. *Jane* switching the gender of her child or *Scott* changing the ages of his children or the brand of pain medication he would use.
In a few cases, *Scott* would get carried away with ad-libbing: When describing his pain level as "a ten on a scale of one to ten" (GP10), causing disbelief in the GP, or when adding having a rash on his chest to his symptoms (GP11).

In two cases, mishaps would occur, irritating the GPs. *Scott* would, in one consultation, give different information on his age (GP10), which the GP would later mention as irritating in the walkthrough. *Jane* would get the gender of her child mixed up (GP13).

The printed-out photo of a sore throat/of swollen and reddened fingers sometimes caused irritation as well. In one consultation, the sore throat depiction *Scott* would hand over to the GP would be commented on with "*This isn't your throat, the person has a beard and you don't*" (GP12).

GP: When you've seen any GPs in the past, have they said what they thought it might be?
S: I can't remember now.. it was a long time ago, to be honest.
GP: And did the antibiotics make any difference or did it still last as long whether you'd have them or not?

S: Well actually I haven't taken them now for a long, long time, but when initially I did get it, it did clear it up actually, but, erm, the thing is, it keeps coming back.
GP: Ok. So.. what did you come in today expecting?

Quote 5: Realistic vagueness. Scott with GP8.

Especially with *Jane*, who tended to gesticulate with her hands and fingers to underline her pain, directing attentive gazes of the GPs to them (e.g. GP1, fig. 4.2, left), the print-out, which was shown typically later-on in the consultation, after a verbal (and gesticulating) explanation of the symptoms had taken place, would be in dissonance with her healthy hands. This was explicitly mentioned by GP10 in the walkthrough interview, who said, he would have diagnosed her earlier, had he seen the depicted sight on her real hands.

The fact that the simulated patients had nothing more to provide when prompted for examination was sometimes mentioned as limiting (GP1, GP6) and an unrealistic course of a consultation (GP9, GP10) during the walkthrough interviews. In the walkthrough interviews, GPs would mention further examination they would have deemed important, like, in *Scott's* case, lymph glands (GP2, GP9, GP10, GP14) and stomach (GP9), and, in *Jane's* case, chest (GP1), lymph glands (GP9) and other joints (GP9, GP15). GP12 would add, that *Jane's* scenario normally would be a *"couch one, in the examination"* (GP12), meaning that she would be examined lying down.

During consultation, GPs would sometimes sublimate examination measures they couldn't do into a question, e.g. "*So if I were to squeeze on them, they would feel painful?*" (GP1 with *Jane*, also fig. 4.2, right). In two cases, the print-out was understood as "only" a visual layer, as GP7 and GP14 would perform further examination, possibly in light of possible neurological causes, on *Jane*'s hands after having seen the print-out: Hand grip strength, joint movement and touch feeling sensation. Such situations would be met with the "strategic" vagueness by the actors as described above.

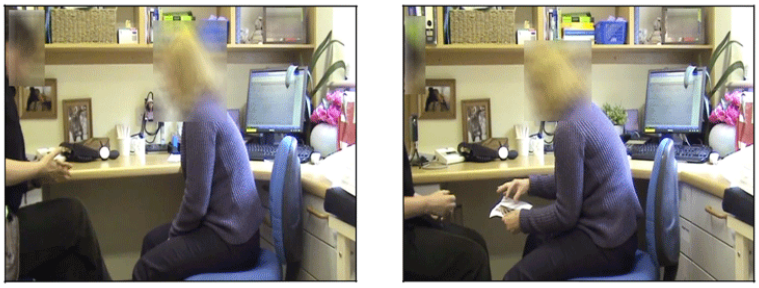


Figure 4.2: GPs attention on the actor's hands as she describes her symptoms (left), Examination with print-out: "What happens when I squeeze there?" (right). GP1.

4.3 Basic Focal Themes in a Consultation

Preliminary remarks

GPs receive specific training on how to structure such medical consultations, as described e.g. in the Calgary-Cambridge guide to the medical interview [43] [44] (see chapter X for further detail). While the analysis of the consultations with *Scott* and *Jane* conducted in the course of this work also hinted on different approaches, attitudes and personality traits present in the attending GPs and influencing their work, the description in this chapter is aimed to avoid an overly analytic view on the medical consultation itself, since this is beyond the scope of this work and would not be above the suspicion of constituting only rather anecdotal evidence.

This chapter leads through the consultations with *Jane* and *Scott* along a set of common focal themes that emerged in discourse analysis of the consultation video recordings. It recounts the course of processing of their cases as observable in the consultations and go into some detail of variability in GP work.

4.3.1 Overview

In this video analysis and its perspective on *focal themes* in consultations, the approach of GPs emerged in four basic themes: (1) The establishment of prevailing symptoms, (2) exploration of personal and family context, (3) the establishment of a construct and

answer to the question of "What is going on?" and (4) the formulation of a plan of action on that basis (fig. 4.3).

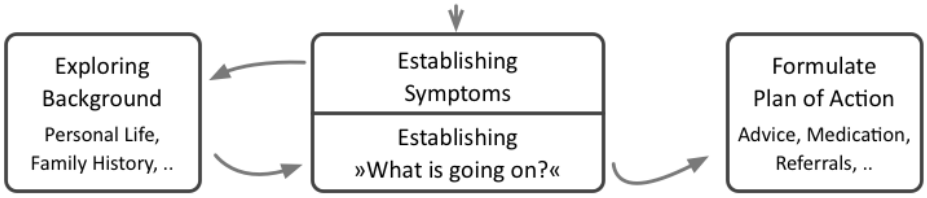


Figure 4.3: Basic Focal Themes in Consultations.

The GPs followed this path in a more or less sequential succession, concluding one such "stage" (e.g. symptoms, background, ..) and then moving on to the next on that basis. Younger GPs (e.g. GP1, GP3, GP5, GP13) would typically adhere more strictly to this process, while older, more experienced ones would be more free in their conversational flow, for example with GP2: He would start the consultation with questions of personal and family background to get to know the personal situation of the patient, and only then get to a first question of prevailing symptoms, approx. 2 minutes in. Instead of, more or less, "finalizing" one stage before moving on to the next, he would iteratively revisit all stages in multiple passes, as he developed his medical opinion, asking more and more fine-grained questions.

Most GPs would periodically summarize and mirror their understanding of the information provided to them by the simulated patients back to them, mostly along the transition from one of these basic focal themes to the next.

4.3.2 Establishing Symptoms

In the course of this stage, GPs would attempt to identify *Jane*'s and *Scott*'s reason for the consultation. This would typically entail (1) questions for the symptoms of the problem at hand, (2) for the patient's ideas (e.g. beliefs regarding the cause), concerns (worries) and expectations (e.g. what help the patient would expect for the problem) and (3) a (simulated, thus limited) physical examination.
(1) and (2) would be met with answers containing the prepared descriptions from the scenarios (see quotes in chapter 4.1 for how they would recite it verbatim). (3) would be met with a paper-printout of a sore throat/swollen and reddened fingers. Most GPs would not ask for further examination, while some would (see section 4.1 for a more detailed description).

Apart from small variability in the approach among GPs, mostly in the amount of time that would be devoted to exploring the patient's view (their ideas, concerns, expectations), this step would vary little among GPs. Sometimes, especially older, more experienced

GPs would weigh in a little on the patient's view (e.g. quote 6), while generally and consistently, the prerogative of interpretation was attributed to the patient in this step.

GP: The problem is, we all get arthritis, and bad stories travel fastest.
J: Yeah.
GP: And the majority who have arthritis, but without it affecting them too much. Those stories don't travel so quickly.
J: Yeah.
GP: So your major concern is, that this is it. You're on top of a slippery slope, you already started sliding down, have given up the exercise classes and what's next?
GP: Yeah.

Quote 6: GP2 relativizes symptoms by experience. He would later judge her symptoms as not too serious.

4.3.3 Exploring Context

After establishing the prevailing problem situation, GPs would widen the focus of the consultation away from it to its history as well as the personal and family context of the patients. Questions would be directed at (1) the symptom history and (2) possible environmental factors. This step would take different courses depending on the scenario:

In *Jane*'s case, the GPs would turn to symptom history and ask only few questions on her personal and family context. *Jane* would mention that, except a short episode of similar symptoms around the birth of her child, this had been the first time she experienced this (quote 2), and often weave in how the joint problems would limit her in her daily routines.

In *Scott*'s case, the GPs would explore his personal situation of stress and how it relates to his symptoms. He would reiterate the descriptions he would often have already given in his opening statements, on how his sore throat "keeps coming back" and his personal situation of commission-based work as a telephone salesperson would prevent him from taking sick leave (see section 4.1 for verbatim quotes).

Apart from this, this stage exhibited little variability across GPs. It however already indicated variability in opinion in both scenarios, that will be discussed in the next stage.

Establishing a contextual background to the symptom situation at hand is an important step in the GP in his process of forming a medical opinion, as context helps disambiguating the underlying reason, separating the common from the uncommon. Results from this step are deemed important by GPs and always documented in the EHR system (for example, quote 7).

I: Anything else come to your mind about why you put those things on?
GP: I think I was trying to convey then that I didn't think he had a known immune deficiency, that there wasn't anything else.

Quote 7: GP3 in walkthrough interview on contextual information.

4.3.4 Construction of 'What is going on?'

Generally, the process of constructing a medical view on the problem at hand is not confined to a single stage in the four-stage-framing of a medical consultation that is used in this chapter, but rather takes place continuously throughout the consultation, advancing around several focal points of interest like the establishment of the medical problem's history and duration or the examination of the patient.
However, once GPs would have gathered and processed enough information to form a medical opinion, they would enter a stage of discussion and joint construction of "What is going on?". This entails discussing ideas on the reason for the problem at hand with the patient. GPs would, in varying degree, involve the patients in their thought process, interweaving the symptoms, concerns and expectations previously uttered by the patients with their developing view. In the Calgary-Cambridge guide, this relates to a recommendation of relating the GP's view to the patient's illness framework in order to achieve a shared understanding. [43, pt. 43]

Jane

In *Jane*'s case, speculation of chronic arthritis would happen in the majority of cases, with two prevalent explanations of cause: (1) "Wear-and-tear"-arthritis and (2) rheumatoid arthritis. GPs would typically remain guarded about giving definitive opinions on diagnosis and stress that further tests are necessary to disambiguate the cause of her problem: *"Certainly with this, Jane Hallam, I don't feel.. that.. it could be a number of things.. it could be RA, [...] in my head I probably know what was going in with it, but I'd want to back that up with further investigations.. and. .I think until you got that.."* (GP1). As they described in the walkthrough interviews, most GPs suspected rheumatoid arthritis to be the underlying cause of *Jane*'s problem (see also figure 4.5 for an overview of consultation outcomes).

Only GP2, while still issuing the tests, would judge her to be too young for having arthritis and suggest other causes.

Scott

In *Scott*'s case, GPs would consistently interpret his sore throat as an infection and attribute it to a viral cause, except 2 GPs who interpreted it as bacterial (GP11, GP14). The majority of GPs (GP1, GP2, GP3, GP5, GP8, GP10, GP12, GP13, GP14, GP15) would relate it to his environmental stress factors.

With *Scott*, variability unfolded in the stage of constructing an idea of "What is going on?" along the question of to what degree his environmental stress situation would be considered as a part of the medical problem at hand, influencing the medical opinion on problem cause and a reasonable further plan of action in the GPs. Depending on this point of view, his symptom situation would, in the extremes, either look like a case of sore throat or a recurring, stressful situation in need of a systemic approach, while the majority of GPs would attribute more importance to his environmental stress. His description of the problem as sore throats that *keep coming back* would be picked up by the majority of GPs as the turning point away from a simple sore throat problem towards a more complex, ambiguous situation: "*Well I mean if somebody is persistently feeling unwell, you have to think about countless problems*" (GP1).

"This particular case had a week history of sore throat, which is a minor problem, but to him it was a major problem, because he was having recurring sore throats, and this was having a knock-on effect and financial implications on his life in general, it wasn't just his sore throat, it was the impact of his sore throat, so it became more than a minor problem, because it became fairly significant for him. [...] If this problem continued, he was going to psychologically.. was having quite an impact at the moment, and so I thought we should.. because he said that he is feeling very stressed, three children, mortgage, main bread winner, he doesn't get paid when he's off work, and these episodes can last up to three weeks." (GP14)

GP: You can take painkillers. We can give antibiotics, but the evidence says it doesn't make much difference. And it's your choice, really. If you want antibiotics, we would prescribe them.
S: But if you don't think it's gonna do anything ..
GP: .. I .. I tried to discourage you, if possible, from having them. And yes, if you'd like anything stronger for the pain, that would be .. we could..
S: No, I mean I appreciate there's no point in taking stuff that's not gonna work, but I mean, .. I'm just open to any suggestions you might have, I'm handing it over to you, I don't know what you'd suggest ..
GP: Well my suggestion is that you understand the situation, which is that you caught another virus this time, which will get better, and it's up to you what you do in terms of pain relief ..
S: I was just sort'a hoping for some way to sort of knock this on the head, but you're telling me that's not possible, there's no way to prevent this.
GP: The evidence would suggest it's better to just leave it alone than to have antibiotics.
S: Hm. Okay.

Quote 8: GP10 demonstrates his stance on "illness vs. disease".

The GPs' stance on this would be observable in the consultations - e.g. in GP2, who would give comparably more space to establishing a clear picture of influencing factors and personal surroundings of the simulated patients and more weight to their views and opinions, or in GP10, who would demonstrate a disease-centered view in his consultations,

especially with *Scott* (quote 8). Similar views were transported by GP11 and GP12, who saw *Scott*'s sore throat as "normal" for the then present winter months.

4.3.5 Formulation of a Plan

In relation to the found understanding of the medical problem, GPs would formulate a plan of action and discuss further steps with the patients. This would, typically and depending on the scenario, entail the prescription of medication as well as further tests and, in two cases, referrals to consultants. All GPs would in varying degree attempt to involve the patients in this stage and suggest action rather than be authoritative. This again would be consistent with training, e.g. as recommended in the Calgary-Cambridge guide as step towards "shared decision making" towards improving subsequent adherence to the decision by the patients. [pt. 48][43].

Scott

With *Scott*, the plan of action varied with the respective problem understanding. GPs with more appreciation/sensibility for *Scott*'s environmental hardship would plan a more extensive regimens of tests and give more advice. They would furthermore shape this advice in a rapport-esque conversation-situation, exploring this hardship together with *Scott*. While all GPs except GP6, GP7 and GP10 did have at least *some* comment on stress handling, many would recite it in a more distanced manner; as a number of recommendations, e.g. to sleep more or reduce stress.

In roughly 2/3 of consultations, a throat swab was done. In 1/3 of cases, a blood test was issued, e.g. by GP4, to check for diabetes. In one case, a referral to an ear-nose-throat (ENT) specialist was issued. The minority of GPs (GP3, GP6, GP7, GP10, GP13) would either prescribe pain medication or recommend over-the-counter versions of it. In some cases, antibiotics were prescribed. Advice was given to *Scott* on dealing with his environmental stress situation, on maintaining a healthy diet and the importance of hand hygiene. Further recommendation included over-the-counter vitamin supplements and to gargle with salt water. GPs would generally remain with him on the basis that he should focus on handling his stress and return if needed, or when results are in, if tests were issued.

Jane

With *Jane*, the plan of action would exhibit little variability. In line with the problem understanding reached by GPs with her scenario, further tests would be necessary to narrow down the underlying cause of her medical problem, most prominently to examine the possibility of a possible auto-immune disease. All GPs issued a blood test, two GPs also issued a hand x-ray. Some GPs indicated, that a referral to a rheumatologist will be issued in subsequent consultations and with blood test results in. All GPs tasked *Jane* to return, when the blood test results would be in with some variability in perceived urgency. GP10 advised *Jane* to call him in 24 hours for the results.

2/3 of GPs furthermore prescribed anti-inflammatory medication, some also prescribed pain medication or recommended over-the-counter versions of it. Further recommendations included over-the-counter glucosamine tablets, wearing a wrist splint at night. Some GPs also gave an information leaflet on rheumatoid arthritis to *Jane*.

Advice and recommendations would vary across GPs and also depend more on the individual situation with *Jane* and *Scott* in the consultations. GP1, for example, reacted to *Jane's* questions for help with her situation with recommending over-the-counter glucosamine tablets "*because she asked me if there's anything **I** can do*" (GP1), apparently making a difference between what she did 'officially' in reaction to the medical problem at hand (issue a blood-test) and what she personally could furthermore do.

Jane's case had triggered at least some suspicions of rheumatoid arthritis in all GPs. Variability unfolded in how urgent her case would be perceived, which was observable mostly in the stage of plan formulation and discussion. In the interaction with the patient, this urgency would be observable in how GPs discussed with the patients the importance of "speeding up the system" or "not delaying them". GP1, when asked whether any guidelines would affect her work, answered: "*There are guidelines, which. erm.. were saying anybody with possible rheumatoid needs to be referred to a specialist within 12 weeks of their problem, so that's what was alerting about her.. she says a few months.. she's almost into her .. so that's why I sent her straight off.. get your bloods, come straight back, and then I'll refer you straight to the rheumatologist*" (GP1). Also GP3, and several other GPs, would describe ringing alarm bells: *"[I had a diagnosis] probably in the first minute.. I think early.. it rang alarm-bells. She's fit and well.. She had 6 weeks of joint swelling and stiffness.. rang alarm-bells. This needs to be investigated, this isn't just a reassurance thing"* (GP3).

Further variability with *Jane*'s scenario emerged in the blood-test issued by the GPs: All GPs but GP1 and GP2 would issue a specific blood-test to determine a possible auto-immune disease as underlying cause for *Jane's* symptoms: the *rheumatoid factor* (RF). With surgery HS, this test was not issued by two of the three GPs of surgery HS (GP1, GP3). In her walkthrough interview, GP1 explained this with specific preferences of the resident rheumatologist in the area of surgery HS: *"Our rheumatologist doesn't want the rheumatoid factor.. he thinks it's a waste of time. We just do normal blood tests, and when they come back as abnormal we'll be referring"* (GP1). Apart from GP1 and GP3 of surgery HS, all GPs issued RF tests.

4.4 Overview of Consultation Outcomes

The following tables (4.4, 4.5) provide an overview of the problem explanation given to *Scott* and *Jane* by the GPs during their consultations as well as prescription of medication, issued tests and further advice and recommendations given by them. Furthermore, in *italics*, the figures give some characterizing words to each consultation, also encompassing how the further plan of action is communicated by the GP.

These characteristics do not exhaustively summarize the consultations, but rather should give orientation.

GP		Prescription	Tests	Other recommendations and advice
GP1	Surgery HS / EMIS LV	*"We have to treat" the **infection,** and Scott should reduce stress. Surgery will contact Scott to consult if throat swab results indicate further action.*		
		Antibiotics	Throat swab	*Stress coping advice*
GP2		*Scott is a "poor fellow", who suffers from an **infection** and **fatigue.** He should reduce stress. Scott is instructed come back for throat-swab results.*		
		Antibiotics, Pain meds	Throat swab, Blood test	Stress coping advice
GP3		*Probably **viral infection**. Next time (with the next infection or if swab results indicate it), blood test will be done and pain medication prescribed.*		
		none	Throat swab	Stress coping advice, gargle with salt water
GP4	Surgery OS / SystmOne	***Viral infection**. Blood test to check for diabetes. Surgery will contact Scott via SMS if tests indicate further action.*		
		none	Throat swab, Blood test	OTC pain medication
GP5		*Probably **glandular fever**. Scott should talk to his boss about his workload. Scott instructed to schedule blood test at the surgery's front desk and come back when results are in.*		
		Pain meds	Blood test	OTC vitamin supplements
GP6		*Viral **infection.** Infections in the frequency of his past are probably normal. He is instructed to come back if things progressed or he is worried.*		
		none	*none*	OTC pain medication, dietary advice, maintain not smoking
GP7		*Viral **infection.** Scott is instructed to come back if he has a further episode. Antibiotics, throat swab next time, possibly referral to ENT specialist.*		
		none	*none*	OTC pain medication
GP8	Surgery EH / EMIS LV	***Infection.** Scott should reduce stress. GP will speak on the phone with Scott about throat swab results, whether they warrant change in medication.*		
		Pain meds	Throat swab	Stress coping advice
GP9		*Stress is not sufficient to explain his symptoms. **Further investigation** is necessary. Scott will be contacted by GP. If he has not heard from him in ten days, he should call the surgery.*		
		Pain meds	Throat swab, Blood test	*none*
GP10		*Scott should put a brave face on it and understand the situation that he has a viral **infection** and 'it will get better'. Scout should consult again if the problem is not settling or he wants stronger pain medication.*		
		none	*none*	OTC pain medication
GP11	Surgery MS / Vision	*Probably bacterial **infection.** This is normal, since 'it is that time of the year'. Scott should sustain a good hand hygiene and use mouth wash. If the problem persists, Scott should come back for a referral to an ENT specialist.*		
		Antibiotics	*none*	Hand hygiene advice
GP12		*This **infection** is normal for winter months and his stressful lifestyle. Scott is instructed to come back in 5 days for results, or if condition worsens.*		
		none	Throat swab, Blood test	Dietary advice
GP13		*Viral **infection.** Scott should continue using his over-the-counter pain medication. He is inistructed to come back, if stress doesn't get better, for a depression-screening.*		
		none	*none*	Dietary advice, stress coping advice, OTC pain medication
GP14	Surgery SE / EMIS LV	*Probably bacterial **infection.** Scott is instructed to return if antibiotics don't alleviate problem in 1-2 days.*		
		Antibiotics	Throat swab, ENT specialist referral	Gargle with salt water, OTC pain medication
GP15		***Infection.** Scott is instructed to come back for blood test results, and generally to come sooner, when he has the next episode.*		
		none	Throat swab, Blood test	Gargle with salt water

Figure 4.4: Overview of consultation outcomes with *Scott* / The Sore Throat scenario.

4.5 Summary and Conclusion

This chapter described the two simulated cases presented to the 15 GPs and how they were enacted as well as the how the consultations with *Jane* and *Scott* progressed and what outcomes they yielded. The main results described in this chapter are described in the following section on *key take-aways*.

GP		Prescription	Tests	Other recommendations and advice
GP1	Surgery HS / EMIS LV	*Could be wear-and-tear **arthritis** or rheumatoid arthritis. If former, then possibly pain medication on regular basis, if latter referral to rheumatologist. GP1 doesn't want to delay Jane with physio or other hypotheses.*		
		none	Blood test	*OTC glucosamine tablets*
GP2		*Jane is a bit young to be attacked by arthritis so vigorously, so GP2 recommends to open it up a bit. Could be **gout.***		
		Anti-Inflammatory meds	Blood test, Hand X-ray	Wear a wrist-splint at night
GP3		*Could be wear-and-tear **arthritis** or rheumatoid arthritis. Jane is instructed to come back for blood test results. Maybe referral to rheumatologist then.*		
		Anti-Inflammatory meds	Blood test	none
GP4	Surgery OS / SystmOne	*Could be wear-and-tear **arthritis** or rheumatoid arthritis. GP4 gave Jane a RA leaflet and will call Jane to discuss blood test results and further steps.*		
		Anti-Inflammatory meds	Blood test	Leaflet given on Rheumatoid Arthritis
GP5		*Could be wear-and-tear **arthritis** or rheumatoid arthritis. Jane is instructed to come back as soon as blood test results are in; GP5 counted the days this would normally take.*		
		Anti-Inflammatory meds, Pain meds	Blood test	Leaflet given on Rheumatoid Arthritis
GP6		*Probably **rheumatoid arthritis.** Jane is instructed to make appointment at front desk. Surgery will call her when results are in, then possibly rheumatologist and/or x-ray.*		
		Anti-Inflammatory meds	*Blood test*	none
GP7		*Could be wear-and-tear **arthritis** or rheumatoid arthritis. Referral to rheumatologist depending on blood test results.*		
		Anti-Inflammatory meds	Blood test	Leaflet given on Rheumatoid Arthritis, OTC pain medication
GP8	Surgery EH / EMIS LV	*Probably rheumatoid **arthritis.** With blood test results in, Jane will be referred to a rheumatologist. GP8 wants to speed things along, since if there's any chance it's rheumatoid, 'they like to see people early in the rheumatoid clinic'.*		
		none	Blood test	OTC pain medication
GP9		*Could be wear-and-tear **arthritis** or rheumatoid arthritis. Blood test results determine further steps, all other issues (ergonomic, wrist band) 'parked for the time being'. Blood test issued on paper because 'that's faster'. Jane instructed to call surgery if GP9 doesn't call her 'within the next 36 hours'.*		
		none	Blood test	OTC pain medication
GP10		*Jane has unspecific **arthritis.** She is instructed to come back for blood test results. In the meantime, she should choose a rheumatologist, who will be contacted furtheron with the blood test results.*		
		Anti-Inflammatory meds	Blood test	*none*
GP11	Surgery MS / Vision	*Could be wear-and-tear **arthritis** or rheumatoid arthritis. Jane is instructed to schedule blood test at surgery front desk and come back 'as soon as possible' for blood test results.*		
		Anti-Inflammatory meds	Blood test	*none*
GP12		*Jane has unspecific **arthritis.** She is instructed to come back 'next week' for further steps. Referral to rheumatologist 'as soon as possible'.*		
		none	Blood test	*none*
GP13		*Jane has unspecific **arthritis.** Without waiting for blood test results, she is referred to rheumatologist. She is instructed to come back in 1 week to discuss blood test results.*		
		none	*Blood test, Hand X-ray, Referral to rheumatologist*	*none*
GP14	Surgery SE / EMIS LV	*Could be rheumatoid **arthritis.** Jane should come back with blood test and X-ray results.*		
		Anti-Inflammatory meds	Blood test	OTC pain medication
GP15		*Could be wear-and-tear **arthritis** or rheumatoid arthritis. Jane is instructed to schedule x-ray and blood test at front desk and come back with results.*		
		none	Blood test	*none*

Figure 4.5: Overview of consultation outcomes with *Jane* / The Joint Pain scenario.

Key Take-Aways

1) The simulated patients were, to the most part, enacted consistently and invoked homogeneous medical responses by the GPs. They were consistently described as credible and the resulting consultations as realistic by the GPs.

2) While some variability was observable, the **consultations presented themselves as overall similar in their course and outcomes** across GPs and surgeries.

The analysis hints on experience and prevailing opinions on the side of the GPs, influencing the reasoning process in the construction of an understanding of the problem and the formulation of a plan, causing variability in how the cases of *Jane* and *Scott* were addressed by them. These differences would unfold on a per-case level: *Scott*'s case divided the GPs slightly along the question on to what degree the construction of a problem-understanding would narrowly focus on the observable signs (a sore throat) or

also take his environmental stress factors into account. This would affect what further measures would be prescribed by the GP. However generally, GPs would follow *Scott*'s enactment of putting main focus on his environmental stress rather than the sore throat problem, interpreting his case as more of a social/psychological problem than the effects of an infectious disease. *Jane*'s case divided the GPs along the question on to what degree her case would warrant for "urgency". This would cause some variability in the plan formulation phase, but would mostly play out implicitly: In GPs crafting the plan towards a swift referral to a rheumatologist.

3) Notably, **regional differences can influence the approach GPs** take to cases of suspected rheumatoid arthritis, as would be hinted on by the blood-tests issued for *Jane* in surgery HS and the influence on them by the rheumatologist local to the area of the surgery.

4) Some **further aspects of variability** were observable around the structuring of the consultation, in that younger GPs would rather adhere to a more step-by-step course through the basic focal theme stages described in section 4.3, while older, more experienced GPs would tend to an iterative approach, e.g. switching back and forth between the establishment of the present problem and the exploration of context, refining the questions with each iteration. Furthermore, the approach of GPs would exhibit variabilities in to what degree they would involve the patient in the decision making process.

While it stands to reason, that the observed variabilities in urgency and the stance on "illness vs. disease" are feasible conclusions on the GPs' thought processes, it is also clear, that a closer analytical look on it would need for more and more specific data. The present view mostly highlights GP-patient interaction and the resulting medical plan along basic focal themes, while the thought process of GPs would probably not be restricted to the sequence of the focal themes, as evident e.g. in some walkthrough interview statements: Some GPs were asked, at what point they would have formed an opinion on the medical diagnosis of *Jane*, and almost uniformly answered: "within the first minutes" rheumatoid arthritis.

Relation to the next chapter

The next chapter focuses on the EHR system usage by the GPs, specifically during the two simulated scenarios and generally, as informed by the walkthrough interviews. It traces the GPs consultation results through the formalization in the course of their documentation in the EHR system.

This further analysis was intentionally separated from the analysis alongside the focal themes done in chapter 4, as the correlation of the forms of variability in consultation course and outcome and variability in EHR usage appeared to be limited. From a content perspective it is evident, that GPs would document their medical opinion, showing the stance they took on issues of variability described in this chapter, e.g. on whether GPs

would see *Scott* as suffering from a mere sore throat or a complex psychosocial situation, but such differences would not affect the way the GPs used their EHR systems.

Thus, the variabilities described in this chapters aren't taken very far into chapter 5. They should not "label" the GPs, as there is only a limited view available on their inner reasoning, and, as stated earlier, it was sought to avoid an overly analytic view on the GPs medical opinions. Rather, a focus is set on how GPs would use the EHR systems to document the case, representing another layer of variability in GP work. Semantic variability in consultations could furthermore pose to be a valuable point of reference in an in-depth language analysis of the EHR documentation, as suggested as possible future work in chapter 8.

CHAPTER 5

Jane and Scott in EHR Documentation

This chapter describes how *Jane*' and *Scott*'s consultations were documented in the EHR systems by the GPs. Section 5.1 (*Comparing EMIS LV, SystmOne and Vision*) contains a comparison of the three EHR systems along key aspects. Section 5.2 (*Content Analysis of EHR Documentation*) constitutes a report on a comparative content analysis conducted on the EHR documentation created by the GPs. Section 5.3 (*Summary and Conclusion*) concludes this chapter and formulates key take-aways.

5.1 Comparing EMIS LV, SystmOne and Vision

This section gives a short overview of the interface of EMIS LV, SystmOne and Vision and then moves on to illustrate the construction of EHR documentation with them by looking at key aspects: What content shape and structure is offered to the GP? How do GPs apply Read Codes? While it stands to reason, that these aspects are not sufficient to exhaustively describe these systems and the effect they have on the resulting EHR data, it seems feasible that they constitute main influencing factors on what EHR data will consequently be available to secondary analysis.

Generally, the description of the systems' attributes in this chapter is based on and grounded in the available video recordings of the consultations with the simulated patients and expert interviews with the GPs. In this, it represents *ground truth* in terms of what EHR functions are *really* used during the consultations as well as the value attributed to them by the GPs.

5.1.1 EMIS LV

Notably, the user interface of EMIS LV is of lower resolution than that of SystmOne and Vision, which relates to its long past. In the walkthrough interviews, GPs mentioned it to be *in your blood* (GP4), "*pretty basic, pretty archaic*" (GP2). GP1 would describe it favorable as a whole (*This one's good. It works pretty well. You have errors, but they're just human errors"*) (GP1). EMIS LV features a limited base functionality well known to many GPs; most of them reported to having at least some experience with EMIS LV. "*[With EMIS], you had frustration about what you couldn't do with it.. [with SystmOne] frustration on how you do it.. because you can do everything with it..* " (GP5).

Its interface has a DOS look and feel and is mostly operated by keyboard. Its functions are navigable by keyboard shortcuts, which some GPs mentioned negatively, as "*[they] are used to mouse-clicking*" (GP1). As the small interface resolution limits the amount of displayable information, EMIS LV features several different views of review and documentation and, again in comparison to SystmOne and Vision, requires the GPs to navigate more in order to obtain overviews over their patients.
GPs criticized this on EMIS LV, as for example GP15: "*What I liked about Vamp Vision, you could have anything on the screen. whereas I hate all this flicking between screens [with EMIS LV].. like you're in consultation mode, and you know, then you'll want to go back to their medical records and on to this.. whereas.. Vision seems to have it all more on the screen altogether.. doesn't it.. there's none of this flicking from one screen to the other*" (GP15).
GP15 also described effects of this EHR design: *"'Trying to be mindful of everything.. when you can only see the consultation screen.. people go into it different ways as well.. don't know whether you've picked up on this.. when I call a patient.. I always go into medical records first.. so that I can see .. have a quick glance on their front screen see what their active problems is or significant problems or whatever.. but some people go straight to consultation mode.. and then you go and don't see any of that.' Interviewer: 'You've lost the history'. GP15: 'Yeah! Often then what's in the consultation .. or major diagnoses if it's a patient you're not familiar with.. you often haven't even seen that..' Interviewer: 'I wonder if working with [Vision] gave you awareness of that sort of things and people only ever working on EMIS wouldn't.' GP15: 'Yeah maybe..'"* (GP15).

Patient Journal Overview

Figure 5.1 illustrates the look and feel of EMIS LV: It offers different overview screens that focus on different aspects ((a): Menu Overview, (b): Timetable Overview). Patient overview: ((c), (d)).
EMIS LV displays an abbreviated version of the consultation's EHR documentation as the patient journal (fig. 5.1, (c)): Several similar overview types exist within EMIS LV; this was the predominant patient overview used by the GPs when familiarizing themselves with the patient beforehand the consultation.
EMIS LV furthermore draws on further structuring of a patient's medical history by giving the user the ability to declare a Read Code a 'Problem Title' (fig. 5.1, (d)). For

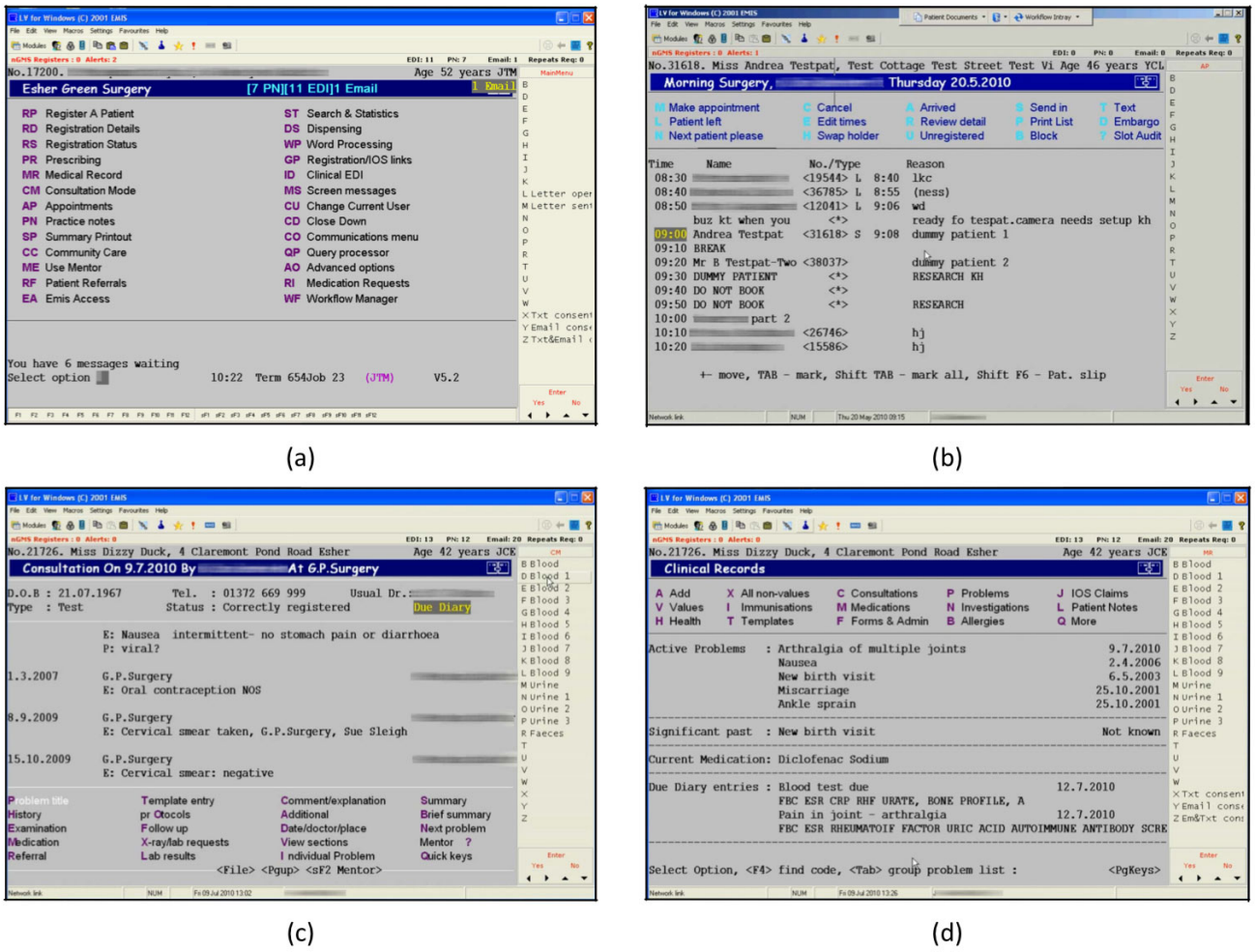


Figure 5.1: The interface of EMIS LV. (a) Overview, (b) GP timetable, different patient overviews (c), (d).

further description of the practice with Problem Titles, see section 6.1.2.

Shape and Structure of EHR documentation

EMIS LV offers its users a variety of options for structuring free text. In usage practice, different approaches to make use of these options are observable (this is examined closer in chapter 5). Some GPs uses little structure, depositing documentation as big chunks within few categories, while others distribute it over more categories. While more specific categories exist, they are hardly used. Usually, generally used categories are **Problem title**, **History**, **Comment** and **Medication**, which are all accessible via keyboard shortcut and listed in a menu in the screen's bottom row (fig. 5.2, left).

Figure 5.2 illustrates GP14 documenting *Scott's* case by entering free-text (left). After completion, the consultation is added to *Scott's* consultation overview (right).

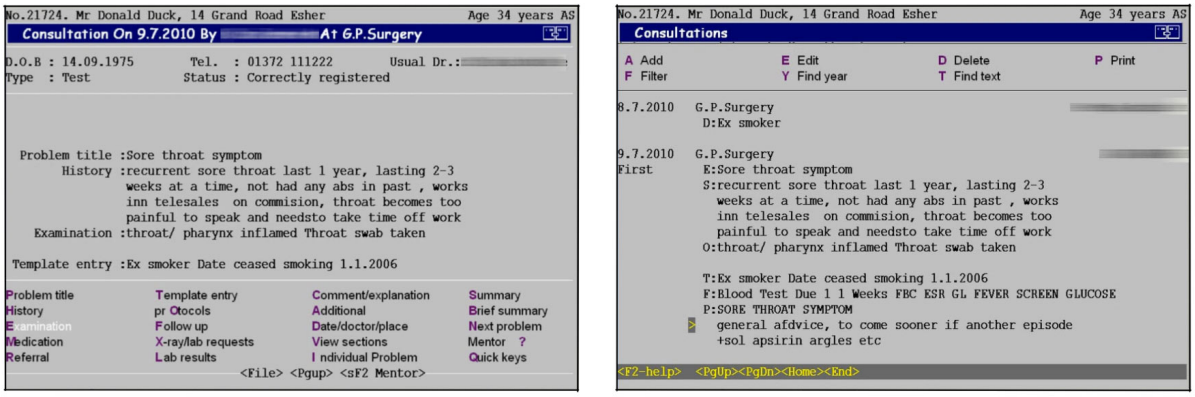


Figure 5.2: GP14 documents *Scott*'s case (who, in this case, was created as "Mr. Donald Duck".

Applying Read Codes

EMIS LV offers categories where documentation can, but doesn't have to be coded, notably **Problem Title** and **Additional**, among others. Figure 5.3 illustrates code-entry: GP14 starts her documentation by adding a **Problem title** to the entry (a), intending to code "Sore throat". The following interface offers her to enter free text, search the Read Code database or search for her codes stepwise by code-category (b). GP14 starts with entering "Sore thr", which gives her a number of options (c). After choosing "Sore throat symptom", a list of more detailed, related entries appears, from which she chooses "Sore throat symptom" (d). EMIS LV offers search results that can be incrementally refined by choice.

The incremental refinement offered by EMIS LV, intended to help the users navigate through Read Codes beyond the space of codes known to them, is described as cumbersome by GP2. As GP2, a GP in his 60s, describes it: "*You can have a shot and just see.. or you might know of a diagnosis.. if that doesn't work, what you can do is.. drill down.. going down the read-code system there.. and for example go into.. it's cumbersome, but you can come down here.. just gently begin drilling down.. interesting how the Read Code system is laid out there.. not sure that that would have got me very far.*".
With *Scott*, GP2 entered "Recurrent Pharyngitis" as the consultation's Problem Title, searching for and coding "Pharyngitis" first. Then wanting to refine this towards "Recurrent Pharyngitis", he just added "Recurrent" as Free-Text, although a fitting Read Code would have been available: *"I was a bit lazy about finding the.. I do find it sometimes difficult to wander around the Read Code system."*

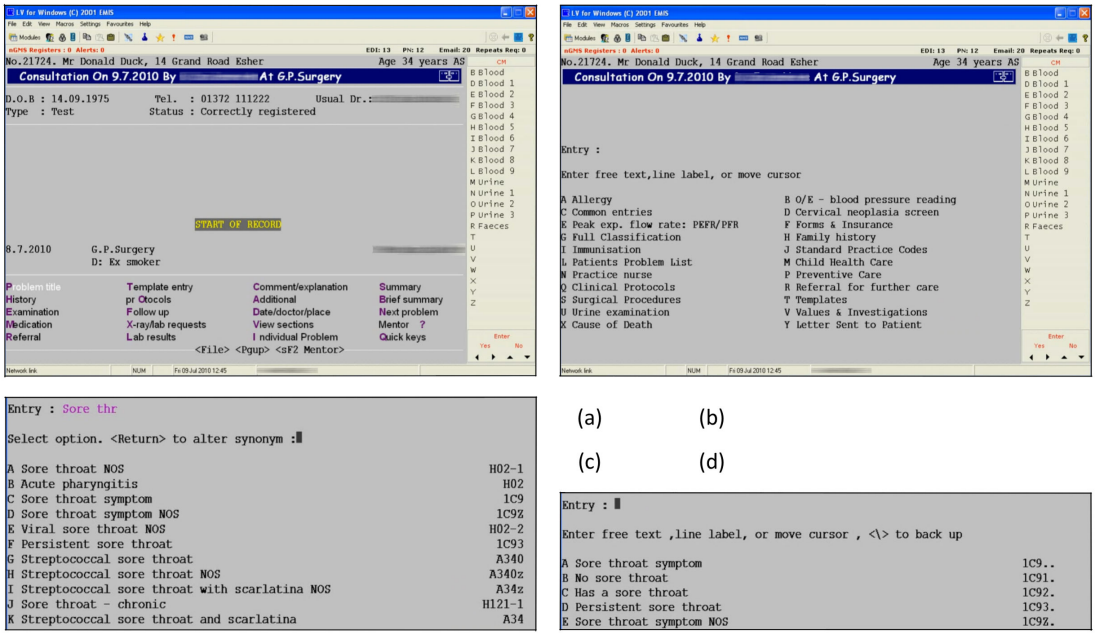


Figure 5.3: GP14 documents *Scott*'s case (who, in this case, is named "Mr. Donald Duck".

5.1.2 SystmOne

SystmOne is the EHR system used by surgery OS. The user interface of SystmOne is not overly cluttered, but densely populated with information. SystmOne is operated by mouse and keyboard and relies heavily on the mouse. *"This is more how you use the computer at home, Internet.., so I think this is a lot more user friendly"* (GP5). Apart from a theme of frustration with SystmOne's predictive coding described in the further course of this section, GPs generally like and prefer it to EMIS LV: *"The power of SystmOne is really exciting.. I'm always discovering things it can do.. I like it for information I can give to patients.. its much more modern. Big fan of SystmOne .. each time I learn something.. I try to bring it into my repertoire"* (GP4).

Patient Journal Overview

While the GPs during consultations with *Jane* and *Scott* mainly used the overview yielded by "New Journal" to familiarize themselves with the patients, SystmOne offers a variety of patient related overview functions, as illustrated in figure 5.4: A Journal dedicated to listing all Read Codes attached to a patient's medical history, a problem overview as well as further detailed lists on medication and vaccinations, among others.
In such detailed overviews, SystmOne would list details that are aspects to consultations (e.g. in the '*Read Code Journal*'), but mostly, a consultation would appear as a whole (fig. 5.4), consultations thus are the predominant structuring element of patient journal

overviews in SystmOne.

In contrast to EMIS LV, SystmOne features a flatter interface hierarchy, requiring fewer steps to switch functions. GP5 mentions this as an advantage in comparison to EMIS LV: "*Easy to dive in and out of things.. you can keep your consultation going.. and you can check their medications.*".

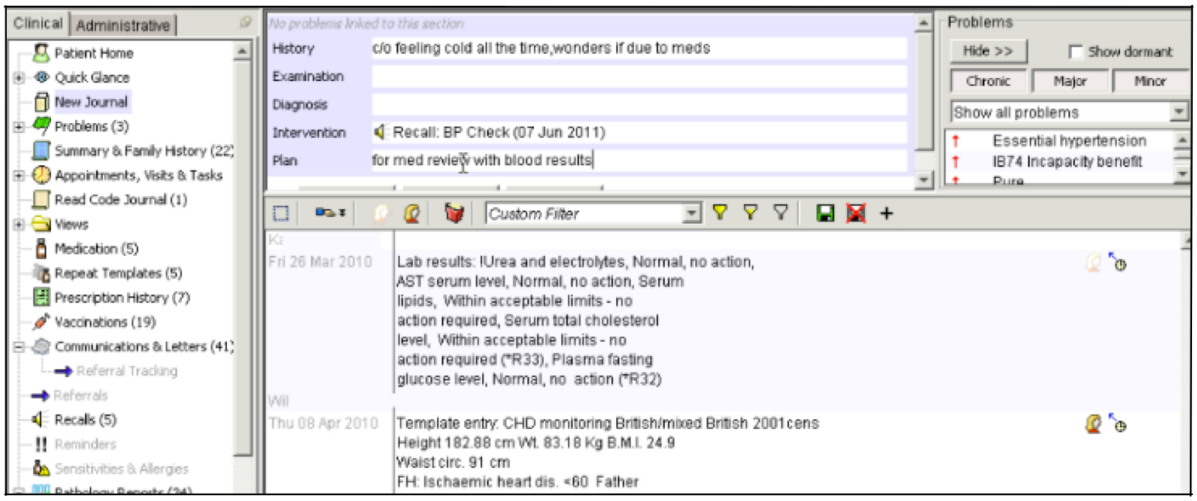


Figure 5.4: Patient journal overview with SystmOne.

SystmOne also offers several features that assist doctor-patient interaction beyond the consultation, e.g. a information service by text message, automated recalls and visualizations of projected medication refill times.

Shape and Structure of EHR documentation

For documenting the main content of a consultation, SystmOne offers a set of five categories: **History**, **Examination**, **Diagnosis**, **Intervention** and *Plan*. Figure 5.5 illustrates how the data-entry-form (top) receives content (middle) and how it subsequently gets added to the patient's consultation journal (bottom). SystmOne disambiguates different types of content by color-coding and icons, e.g. medication and pathology tests, which in SystmOne are furthermore generalized and listed as interventions. Color-coding (green) would also highlight applied codes (fig. 5.6, (d)).

Applying Read Codes

SystmOne promotes the usage of Read Codes by continuously interpreting the text entered into all consultation-related entry fields for matching Read Codes and other text bits and offering them in context menus; it *"automatically comes up with a list of potentials for you"* (GP7). Read Codes can also be searched in catalogues, but predictive choice is the main usage scenario among GPs.

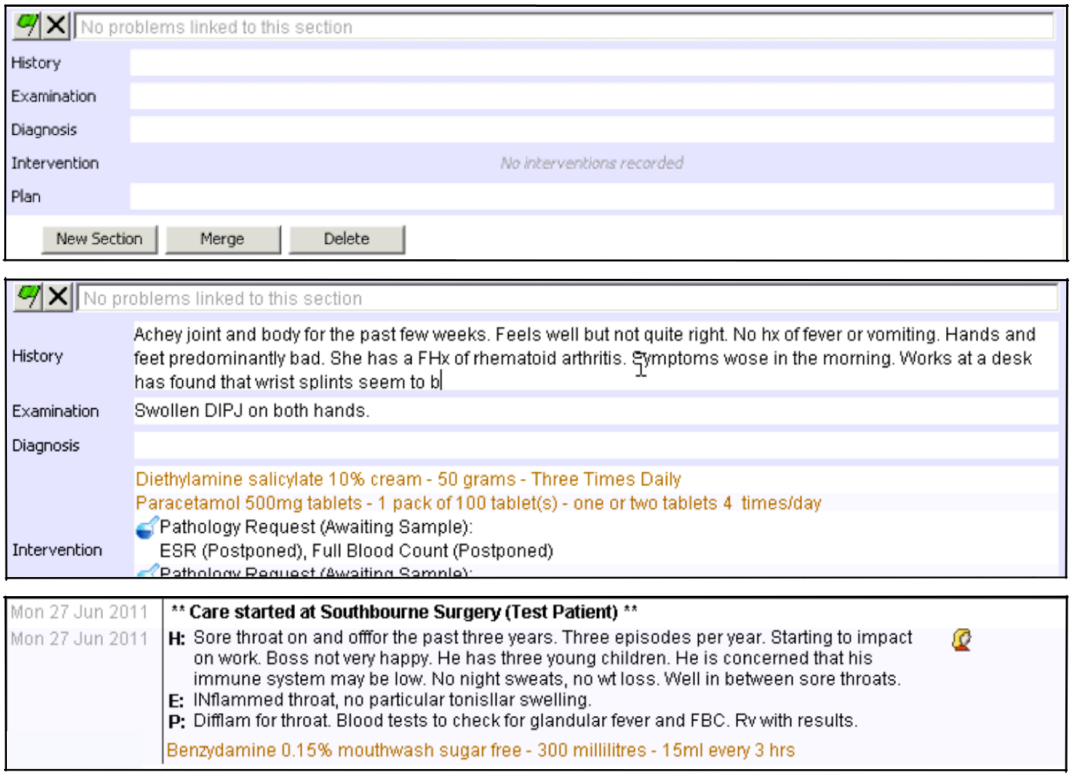


Figure 5.5: SystmOne consultation data entry interface form (top), being filled out (middle) and how it is listed in the patient's consultation journal (bottom).

This design was observed to exert a distracting influence on the case documentation, as illustrated by GP6 documenting *Scott*'s case (figure 5.6): The prediction logic would offer a limited choice of codes at first, which expanded upon the user hovering over it with the mouse pointer. GPs had to actively dismiss the lists for them to vanish. The codes were mostly unwanted by the GP and irrelevant to the case at hand. For example, as GP6 tried to convey that *Scott* had "tried cough medication", the prediction logic would suggest a list of several codes deemed possibly relevant, containing the medical problem "cough" as well as text bits for all kinds of information, like "countries planned to visit" or "court case pending" - a list of mostly irrelevant content that had to be actively dismissed (fig. 5.6, (c)). Such lists would pop up frequently, for further example suggesting "Formula milk fed" or "Forgetful" when GP6 tried to convey "Ex-Smoker *for* 20 years".

In the walkthrough interviews, several GPs recounted frustration with the way SystmOne offers predicted Read Codes:
"With SystmOne, I'm struggling finding Read Codes for. It seems to come up with the most ridiculous Read Codes first. So if you put painful for painful hand, you get „painful

orgasm". That's not something I use that often. Like somebody just does it to get you laughing, but it's not" (GP1).
"The thing about this system, which is a little bit helpful or annoying to many ..." (GP4).
"At the moment, I don't have the setting where I'd be getting every read code.. "It can be a pain. It can be useful but it can be a pain. You have to think all the time whether you'd accept or reject that." (GP5)
"I do have [predictive text] set like that [(turned on)], and it drives me insane. I have to be careful.. because sometimes you might write.. no lymphadenopathy.. and it'll code 'On Examination: Lymphadenopathy'. ***which is completely opposite.. so you have to have your wits about..****"* (GP4).
"I don't touch-type, so a lot of what comes up in history is codes I ignore.. because I'm not looking at the screen and I read back on it" (GP4).

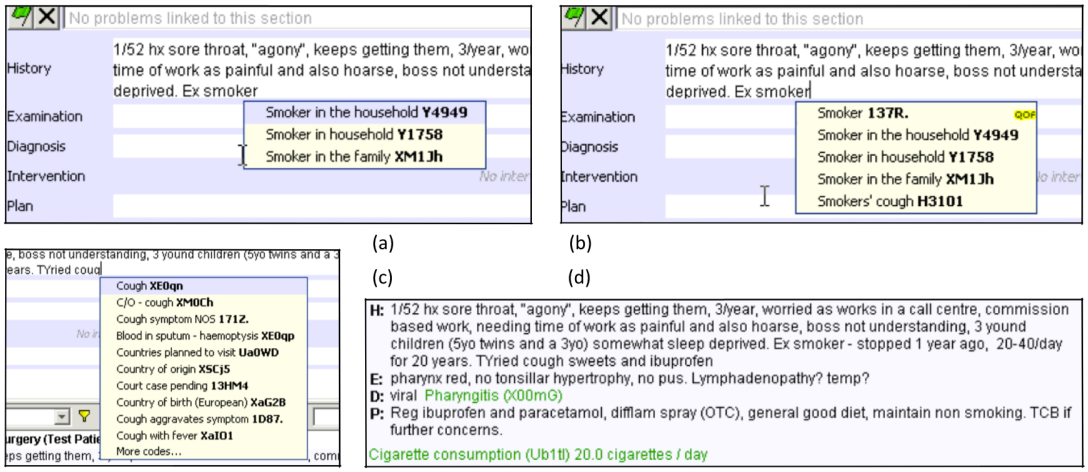


Figure 5.6: Predictive code suggestion in SystmOne.

When documenting *Scott*'s smoking history, GP4 was prompted with a QOF-tagged code of "Smoker" (5.6, (b)), which she subsequently used. This action led her to a multi-purpose template, which GP6 had to navigate through in several steps in order to complete the process. Subsequently and even though a Read Code for "Ex-smoker" exists, the resulting EHR documentation contained the code of "*Cigarette consumption (Ub1tl) 20.0 cigarettes / day*", wrongly attributing the code of "Smoker" and stating that *Scott* would still smoke 20 cigarettes a day. A free text description in the *History* would contain the information as indicated by *Scott*: "*Ex smoker - stopped 1 year ago, 20-40/day for 20 years*" (5.6, (d)).

In the walkthrough interview, GP4 noticed the error: "*I gave him some smoking cessation advice.. I coded that.. I've got a feeling I've coded ex-smoking .. the smoking I've coded because its useful for QOF.. I think I might have made an error here.. it looks like he actually smokes 20 cigs a day now.. - that's what he used to do"* (GP4). (The practice around QOF, the Quality and Outcomes Framework, is described in more detail in

chapter 6.1.) It seems feasible to attribute such errors to the interaction design of this prediction feature, making Read Codes an often hardly relevant nuisance, popping up every couple of words. As mentioned by GPs, this feature can be turned off, but, as practice demonstrates with GP4, this isn't always done properly. In overall comparison with EMIS LV and Vision, EHR documentation created in SystmOne contained the smallest amount of Read Codes, which might relate to this design.

5.1.3 Vision

Patient Journal Overview

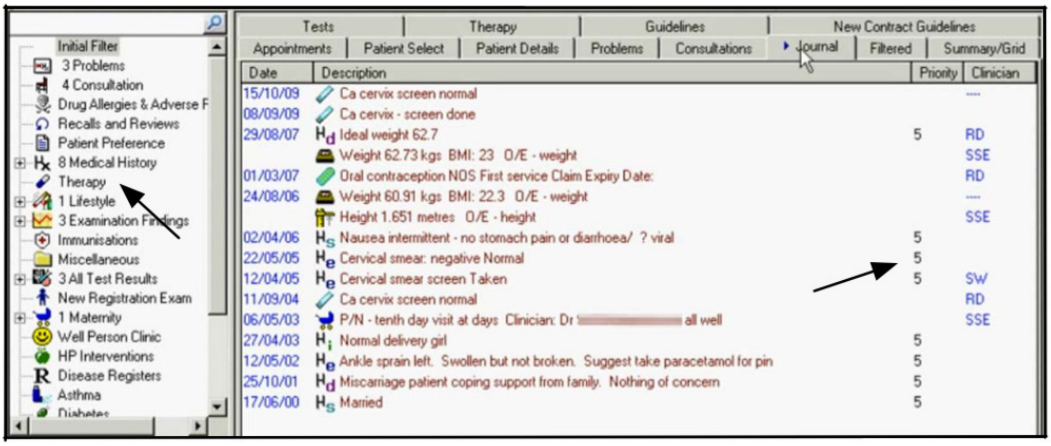


Figure 5.7: Patient Overview with Vision. Vision offers a priority setting in addition to the problem label as structuring elements for summaries (arrows).

As SystmOne, Vision offers a variety of patient-related overview functions in a menu situated at the interface's left in a screen densely populated with information. *"With Vision, you could have anything on the screen"* (GP15).
Unlike the other systems, Vision retains a comparably fine-granular structure when representing the patient's summary and history, as visible in figure 5.7: A patient's journal constitutes a list of elementary entries, that are related to categories and represent aspects/topics of a consultation rather than consultations as a whole.
Feasibly in order to uphold good overseeability also with complex medical histories, as the decision to use these entries as structuring elements, without doubt yields long entries, Vision, in addition to the possibility of structuring a patient's medical history by declaring Read Codes a "Problem Title", offers a priority scale setting as additional structure (fig. 5.7, arrows): Every element can be provided with a priority setting on a ten-level scale (figure: right arrow); elements with a present priority value are counted as medical history elements and add up to the medical history summary count (figure: left arrow).
GP11, GP12, GP13, the GPs who use Vision and whose work was examined in the course

of this thesis, didn't use this function; they rather left it at its standard value setting of '5'. This might imply that this setting is unused, but a conclusion would seem premature, as more fine-grained structure of an EHR entry might only become relevant with more complex medical histories than the two simulated patients came with and would yield during the consultations.

Shape and Structure of EHR documentation / Applying Read Codes

Unlike EMIS LV and SystmOne, the Vision system structures EHR free-text and codes in an integrated fashion, thus a discussion of shape and structure of EHR documentation and applying Read Codes is combined in this section.

As mentioned in the previous section and illustrated in fig. 5.7, a consultation's EHR documentation is structured in different content types not unlike the content categories available to the user with EMIS LV or SystmOne. With Vision, these categories without exception are related to a respective subset of the Read Code list. To use a content category, a code from the corresponding code sub-list has to be entered or chosen. Read Codes thus are the central structuring element of EHR documentation with Vision. The following example illustrates the process:

EHR documentation with Vision starts with entering a Read Code (fig. 5.8). After entering a keyword, the user is offered related Read Codes from which to choose.

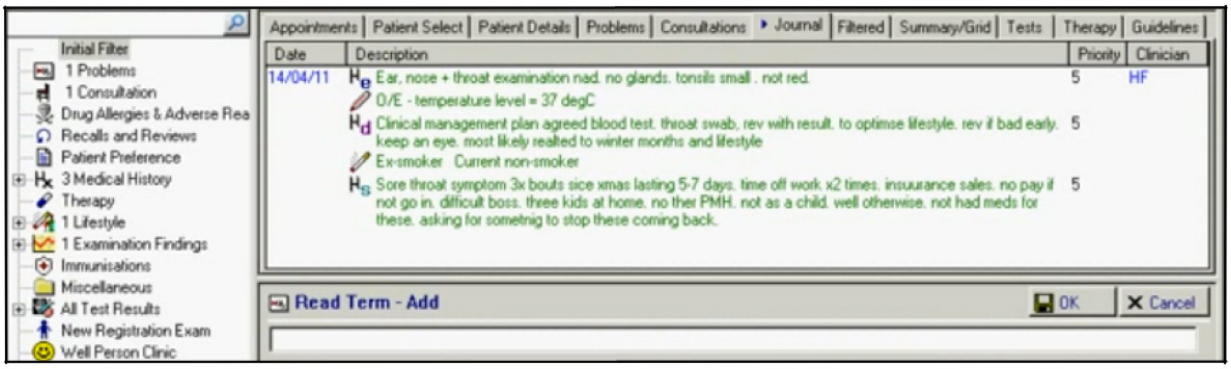


Figure 5.8: Documenting with Vision starts with entering a Read Code.

Depending on the Read code chosen, the content category subsequently to be created is decided. For example: Read Codes denoting a symptom term yield an entry form for a *History/Symptom* category entry, Read Codes denoting a diagnosis term yield one for a *History/Diagnosis* category entry. Read Codes denoting a medical procedure yield a *Procedures, Specimens and Samples* category entry.

When documenting *Jane's* case, GP12 and GP13 documented her family history of arthritis. After typing in and choosing a Read Code out of the 'Family History' subset[1]

[1]Chapter 12 of the Read Code database contains Family History Diagnoses.

(in the case of GP12, '12fZ.00 - FH: Musculo-skeletal dis. NOS'), Vision produces the 'Family History - Add' interface. There, the user can furthermore concretize the entry with free text, if necessary (fig. 5.9).

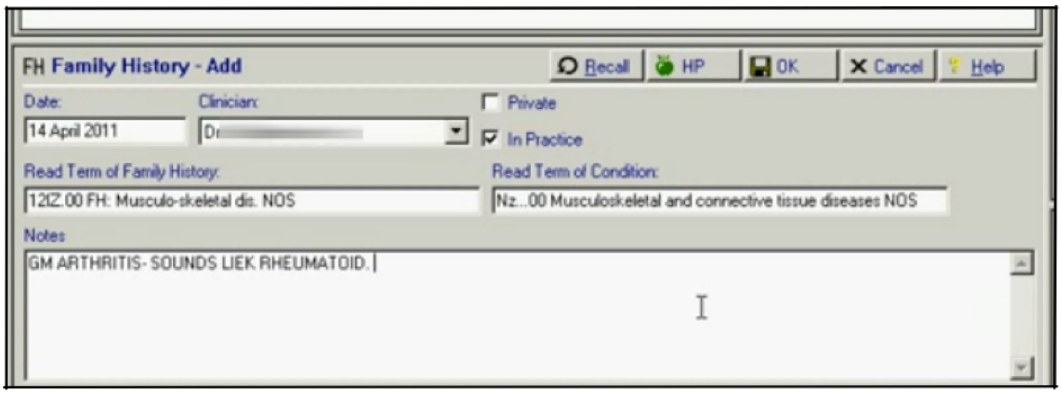


Figure 5.9: Entering Family History with Vision.

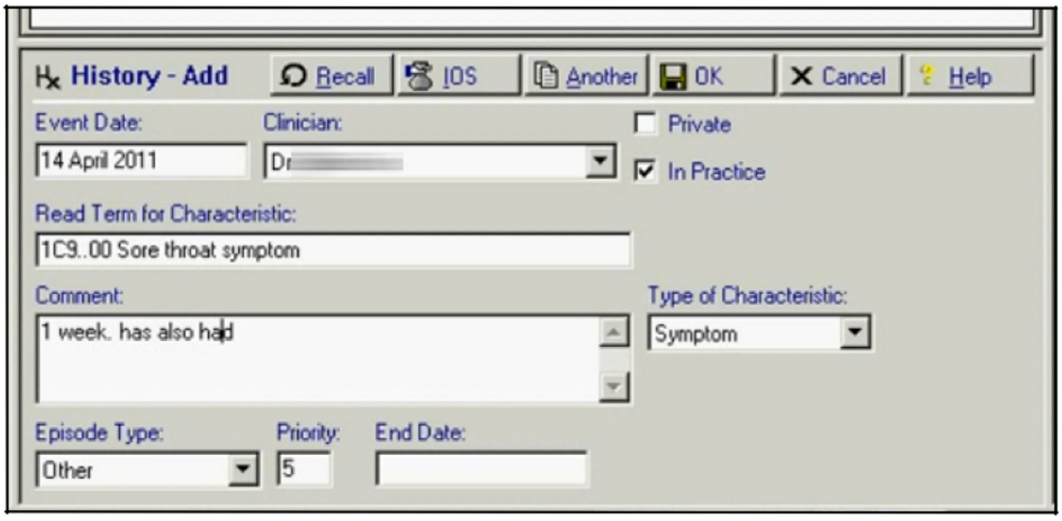


Figure 5.10: A symptom-related Read Code yields the 'History - Add' interface. Vision.

Upon entering a symptom-related Read Code, Vision produces the 'History - Add' interface (fig. 5.10). The GPs mostly used *History/Symptom* and *History/Diagnosis* categories for documenting the consultation. For documenting examination results, e.g. the examination results of *Scott's* throat or *Jane's* hand, History/Examination was used most. GPs would use the comment-field extensively for the major part of their free-text entry. In order to specify a blood test, GPs would choose a related Read Code, which brings up the 'Procedures, Specimens and Samples - Add' interface. The free-text entry

field was used to specify the blood test (fig. 5.11).
Generally, this Read Code-centered organization is used to efficiently trigger relevant specialized entry masks for a wide variety of categories, like medication prescription, smoking status or occupation information, to mention further examples. With Vision, free text is always structured around a Read Code, it is "*gonna be relating to whatever is coded*" (GP11).

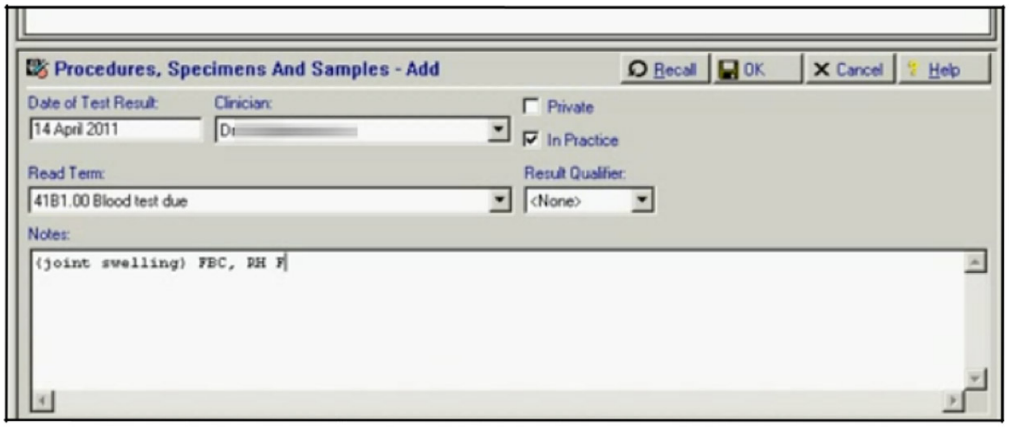


Figure 5.11: A procedure-related Read Code yields a specific, test-related interface. Vision.

This organization also brings it's disadvantage: Users necessarily have to find an appropriate code to enter information, which is a learning process: "*There's a huge spectrum [in code amount across GPs]. You code more as you get more familiar with the system*" (GP11). "*You try and find the code for something and you just can't find what you want, and you have to compromise, and that's a pain*" (GP12). This would affect the EHR documentation in practice, for example with GP11, who, because she would not remember the Read Code for placing examination results, free-texted them: "*There might be [a code] for the examination.. but it didn't come to mind, so I just put it just in the free-text, the examination*" (GP11).

Vision offers a category specifically for documenting family history elements of a patient (fig. 5.9), but demands the user to hinge this information on a Read Code. To GP11, this would prove to be an obstacle for documenting *Jane*'s family history; while generally intending to 'properly' categorizing this information as family history, she would not feel confident to attach a Read Code to the illness of *Jane's* grandmother without ever having seen her: "*I could have coded the family history.. but then you didn't know exactly, what her grandmother had.. what sort of arthritis*" (GP11).

Having overcome this obstacle of just having to know relevant Read Codes, GPs described Vision and this approach to structuring information as efficient. Notably, there is a learning curve towards this skill that arguably causes variability in EHR documentation

practice: "*I'm used to that system now, so I know what to write. [..] I remember the beginning when I was starting.. it was a bit difficult to put in.. erm.. some of the things.. but now I what I'm looking for, so..*" (GP13). "*I've changed the way I use vision, as I got to know how to use it. When I was first here, I would very rarely code an examination or a temperature or something like that, that would all go in the free text, because that was just what was intuitively easy to do, but actually, it's probably quicker to do it, and I do that now, so blood pressures and temperatures all get coded. In examination, you can just enter 'ENT ex' and then you get a box up for ENT examination and you can put that in*" (GP11).

Figure 5.12 shows the finished EHR consultation entry for *Scott's* case as made by GP12. The chosen content categories are signified by an icon. "He/Hd/Hs" denote a History/Examination, /Diagnosis or /Symptoms-category, the pencil-icon denotes a test-result, the cigarette-icon denotes smoking status information.

Figure 5.12: A finished EHR consultation entry of *Scott's* case (GP12). Yellow highlighting signifies codes. Vision.

The respectively used Read Codes are highlighted, illustrating exemplary coding practice with Vision: In comparison with EMIS LV and SystmOne, Vision invokes a diverse code usage on an often generally conceptual level (involving Read codes like "Clinical management plan agreed") in addition to "classic" symptom- or disease-related coding.

5.1.4 Conclusions on System Comparison

In comparing the three EHR systems employed by the surgeries HS, OS, EH, MS and SE - EMIS LV, SystmOne and Vision, differences in how EHR documentation is structured as well as how Read Codes are integrated in this process were observable.

The **user interface** of EMIS LV yields a information in lower density than the interfaces of SystmOne and Vision, which was mentioned by GPs as an obstacle, as it prolongs the user interaction necessary to retrieve information. SystmOne and Vision, both systems belonging to a *next generation* in comparison to the *archaic* (GP2) EMIS LV, furthermore offer a more diverse set of functions like sending text messages to patients via SMS or access to information leaflets.

The three systems have different takes on how to integrate **Read Codes** in their EHR documentation. All offer roughly similar options for structuring information, like fields designated for *examination-related* or *history-related* information, however their integration with Read Codes differs:
While EMIS LV only offers GPs to incrementally search through the Read Code database, SystmOne features "predictive coding": it continuously searches the entered free text for opportunities to add a code, suggesting them in contextual menus. The design of this feature, especially the fact, that those suggestions have to be actively dismissed, might make it difficult to use, e.g. for GPs who are not good at touch typing, as described by GP4, and is consistently mentioned by GPs as a source of frustration.
Vision uses Read Codes as structuring elements of the EHR documentation, both within a consultation and for overviews of a patient's medical history. Free text is not put in independently, like with EMIS LV and SystmOne, but can only be entered in relation to a specific Read Code. Vision requires of GPs to have a certain vocabulary of Read Codes, also ones that would not come up with EMIS LV or SystmOne, in order for GPs to be able to use it. GPs like the efficiency of it, once they have acquired a certain level of skill with Vision, but also indicate, that information, without them knowing the specific Read Code or feeling confident to apply it, would not be structured.

Furthermore, the three systems are based on different concepts of how EHR documentation is structured in general. While basic **categories** are similarly present with all three systems, e.g. all feature categories related to the patient's medical history, examination results or the plan of action, their design differs: With EMIS LV, categories can be chosen from a menu (fig). With SystmOne, categories are present from the start (e.g. visible in fig. 5.4) which might suggest to the user that they need to be filled out. With Vision, categories are tied to Read Codes. As the next section will show, this affects the resulting EHR documentation.

As was illustrated in this system comparison, the different ways EHR systems are designed demand different skills and approaches from GPs for documenting medical consultations. The next chapter will take a closer look on the resulting EHR documentation and further trace variability caused by this design decisions.

5.2 Content Analysis of EHR Documentation

Taking the data structure that is available to the GPs in the brand of EHR system used by them (chapter 5.1) as a starting point and the consultation results (chapter 4.4) as respective points of reference, this chapter reports on a content analysis done on the EHR documentation created by the GPs about their consultations with *Jane* and *Scott*. It focuses on what the GPs documented thematically, how this information was documented (as free text or Read Code) and what content categories were used.

The categories visible in the schematics (figs. 5.13, 5.15, 5.17) do not represent a full spectrum of function available to the GPs in theory, but what structure was used by them in their documentation of the cases at hand. Notably, usage of categories furthermore

needs to be regarded in relation to the respective consultation outcome, e.g. the sparse usage of the "Referral"-category with EMIS LV relates to the fact that, among EMIS LV using GPs, only GP14 would issue a referral.

The figures of the following subsections contain 1) free-text descriptions, denoted by thematic keywords in <..>-signs, 2) coded content, denoted by underlined, thematic keywords enclosed by asterisks: *..* and 3) - on the outer left of the figure - content talked about verbally among patient and GP during the consultation, which is denoted by keywords enclosed in |..|-signs.

5.2.1 EMIS LV

Figue 5.13 depicts an overview of the EHR content documented in EMIS LV by the GPs of surgeries HS, EH and SE.

Category Usage

The GPs of surgeries using EMIS LV - HS, EH and SE - exhibited both similar and divergent usage practice in the categorization of their EHR content in EMIS LV.
The category **History** was used by all GPs and mostly in similar ways: With *Scott*, all GPs used it to document his personal situation and environmental stress factors as well as symptoms/medical problem. With *Jane*, most GPs would document her symptoms/medical problem as well as her grandmother's suffering from arthritis. Three of nine GPs would furthermore use it to document the medication *Jane* had tried in the past. Especially with *Scott*, GPs would devote considerable space in the *History* category for describing his personal situation. Notably, GP9 would accumulate almost all of his EHR documentation in this category, creating one long piece of free-text spanning all sorts of content categories.
The category **Examination** was used consistently for documentation comments resulting from the examination of *Jane's* hand picture and *Scott's* throat picture. Notably and mostly with *Scott*, most GPs would frequently mention "negatives" in respect of the examination, namely the fact that they would not find unhealthy tonsils or lymph nodes.
The category **Comment** exhibited rather diverse usage. Most GPs would use it to document the further plan of action, e.g. with *Jane*, that, with blood test results in, she would come back to discuss further steps. Furthermore, GPs would sometimes explain their reasoning, mention the tests and medication issued or describe further advice and measures recommended by them, e.g. with *Scott* (that he should gargle with salt water, get over-the-counter pain medication or reduce stress). With *Jane*, the *Comment* category would also be the place for GPs to document speculation on their suspicions of rheumatoid arthritis being the underlying cause for her symptoms. GP8 used this category to specify the blood test he issued in *Jane's* case.
The category **Problem title** was filled with keywords denoting a symptom or disease to the consultation. This category was used by most GPs and mostly in similar fashion. Except for GP2, who would specify the Read Code he chose ("Pharyngitis") with a free-texted word ("*Recurrent* Pharyngitis"), this category was used with Read Codes only.

Figure 5.13: Usage of EHR categories in EMIS LV with *Scott's* case (bottom) and *Jane's* case (top) with respective code count (outter right column). <>.. Free-Text, **.. Code, ||.. mentioned in consultation.

The category **Medication** was used consistently for prescribing medication.
The other categories further reflect the diversity in EMIS LV usage across EMIS-LV-using surgeries HS, EH and SE and will be discussed in the following sections on coding and tests/medication.

Coding

GPs using EMIS LV, to the most part, used codes to attach a *Problem title* to the consultation. Further coding encompassed medication prescriptions, blood test prescriptions, as well as - according to walkthrough interviews most likely QOF-related - aspects of the medical history (GPs 1, 8, 9 and 15 would code *Scott's* history of smoking, GP9 furthermore his complexion and alcohol consumption) (fig. 5.13). GP 14 furthermore coded *Jane's* family history of arthritis.

Notably, GP9 of surgery EH described a general agreement among the surgery's GPs on not to use too many codes: *"I think there is a theme in the practice, that we don't want a lot of rubbish data on the problem list.. in other words, codes that are too vague to be useful.. concrete things that affect someone's management in short order.. that's what we're looking at. stuff that's not overly clogged up with vagueness. we are probably similar in that"* (GP9). In comparison with surgeries HS and SE, this statement seems feasible, as GPs from surgery EH would use fewer codes and spread their EHR documentation over fewer categories (fig. 5.13). The theme of surgeries negotiating a common understanding of 'good' record-keeping is further described in chapter 6.3.

Tests and Medication

EMIS LV exhibited variability in how GPs would issue **blood tests**: While GPs from surgery HS used *Pathology Plus*, an EMIS LV add-on, GPs from surgery SE used EMIS LV's own *Follow-Up* category, where they would place the workflow-related Read Code "Blood test due" and specify the blood test's details as free-text. Both ways constitute coding, as *Pathology Plus* also translates blood tests issued with it into a Read Code.

In the case of *Scott*, with whom all GPs from surgeries HS, EH and SE, except GP10, would do a **throat swab**, variability was observable among surgeries: GPs from surgery HS documented it in the *Comment* section, GPs from surgery SE in the *Examination* section. GPs from surgery EH used the *Examination* section (GP8) and *History* section (GP9).

EMIS LV exhibited no variability in how GPs would prescribe **medication**: This would happen consistently in the *Medication* category and be Read-Coded there.

To document **other advice** and recommended measures, GPs would consistently use free text: GP1, GP2, and GP8 would give advice on how to cope with the environmental stress factors affecting him; GP1 and GP2 would document this advice as free text in the *Comment* category, while GP8 left it out. GP3, GP14 and GP15 recommended gargling with salt water. GP3 left this undocumented, while GP14 and GP15 documented it as free text in the *Comment* section. GP10 and GP14 recommended using over-the-counter pain

medication to *Scott*. GP10 left it undocumented, GP14 documented it in the *Comment* section.
With *Jane*, GP8, GP9 and GP14 recommended over-the-counter pain medication, with GP9 documenting it as free-text in the *History* category and GP8 and GP14 leaving it undocumented. Further advice by GP1 (over-the-counter glucosamine tablets) and GP2 (wearing a wrist-splint over night) was documented in the *Comment* section.

Word Count

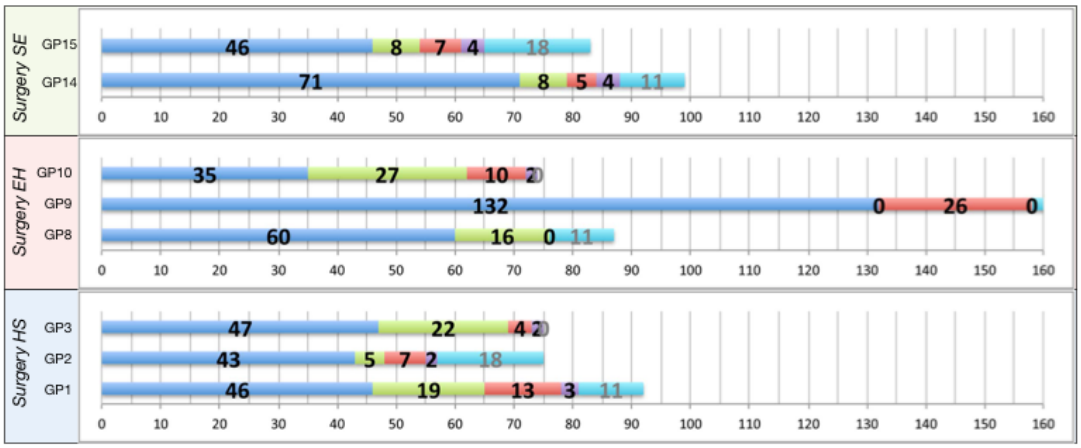
Figure 5.14 illustrates category usage by the GPs of surgeries HS, EH and SE by word count. As hinted on by surgery averages, the length of EHR documentation might be influenced by the habit of the respective surgery. GPs from surgery HS, across scenarios, have the smallest word count, while GPs from surgery EH average notably higher.

5.2.2 SystmOne

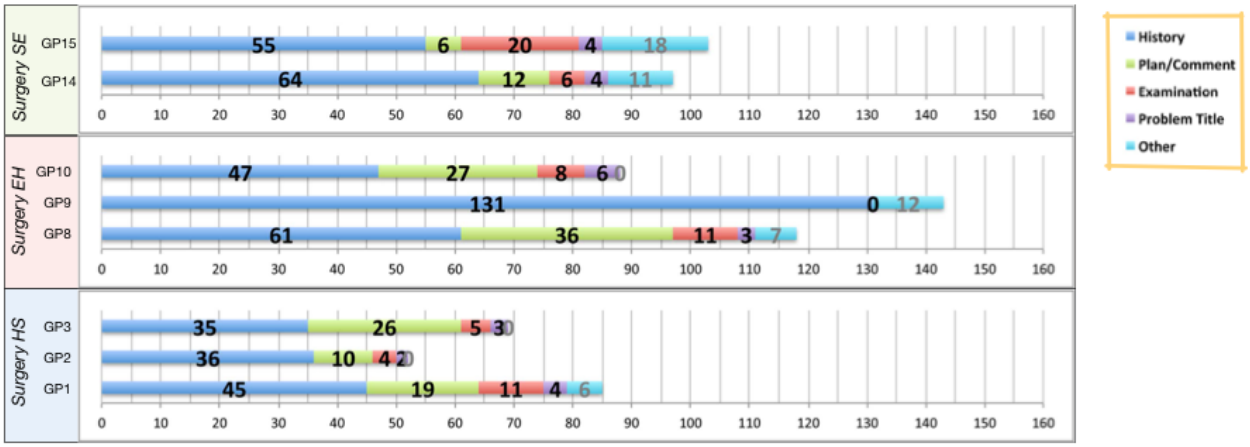
Figure 5.15 yields an overview of category usage with Vision.

Category Usage

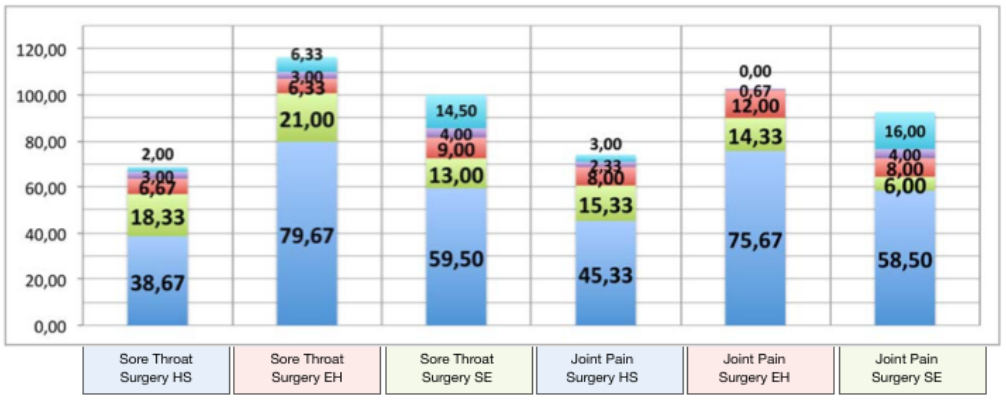
The four GPs of surgery OS, using SystmOne, structured their EHR documentation by the categories visible in figure 5.15. The categories *History*, *Examination* and *Plan* were used by all GPs, while most GPs (*Scott's* case)/only one GP (*Jane's* case) used the category *Diagnosis*. The categories of *Intervention/Medication* and *Intervention/Pathology Request* would be filled out automatically on the basis of templates.
The category **History** was used consistently by the GPs to document the symptom/problem situation of both simulated patients, as well as case-dependent aspects the GPs deemed noteworthy: In *Jane's* case, her description of her memories of her grandmother having been affected by arthritis was documented by 3 of 4 GPs, and in all of those cases under the *History* category. With *Scott*, 3 of 4 GPs would describe his environmental stress situation there. GPs would furthermore use this section to mention medication the patients had tried in the past (JP: GP4, GP7).
Examination was used to describe the results of the examination done by the GPs in both scenarios. GP4, with *Scott*, would free-text a diagnosis there. This category contained many *negative* mentions as well, e.g. "no deformities" (GP5/*Jane*), "no tonsillitis" (GP4/*Scott*) or "no cervical LN" (GP5/*Scott*)).
Plan was used in an "all-purpose" way, containing short summaries of the consultation and the measures taken by the GP, reasoning for the prescribed medication and tests. Recommendations of over-the-counter medication and other advice given during the consultations would be documented there. Especially with *Jane*, GPs would also use this category to document their suspicions of her symptoms being caused by rheumatoid arthritis.



(a) Joint Pain (JP): Category word count in EMIS LV



(b) Sore Throat (ST): Category word count in EMIS LV



(c) Average category word count in EMIS LV

Figure 5.14: Word count in EHR documentation within EMIS LV.

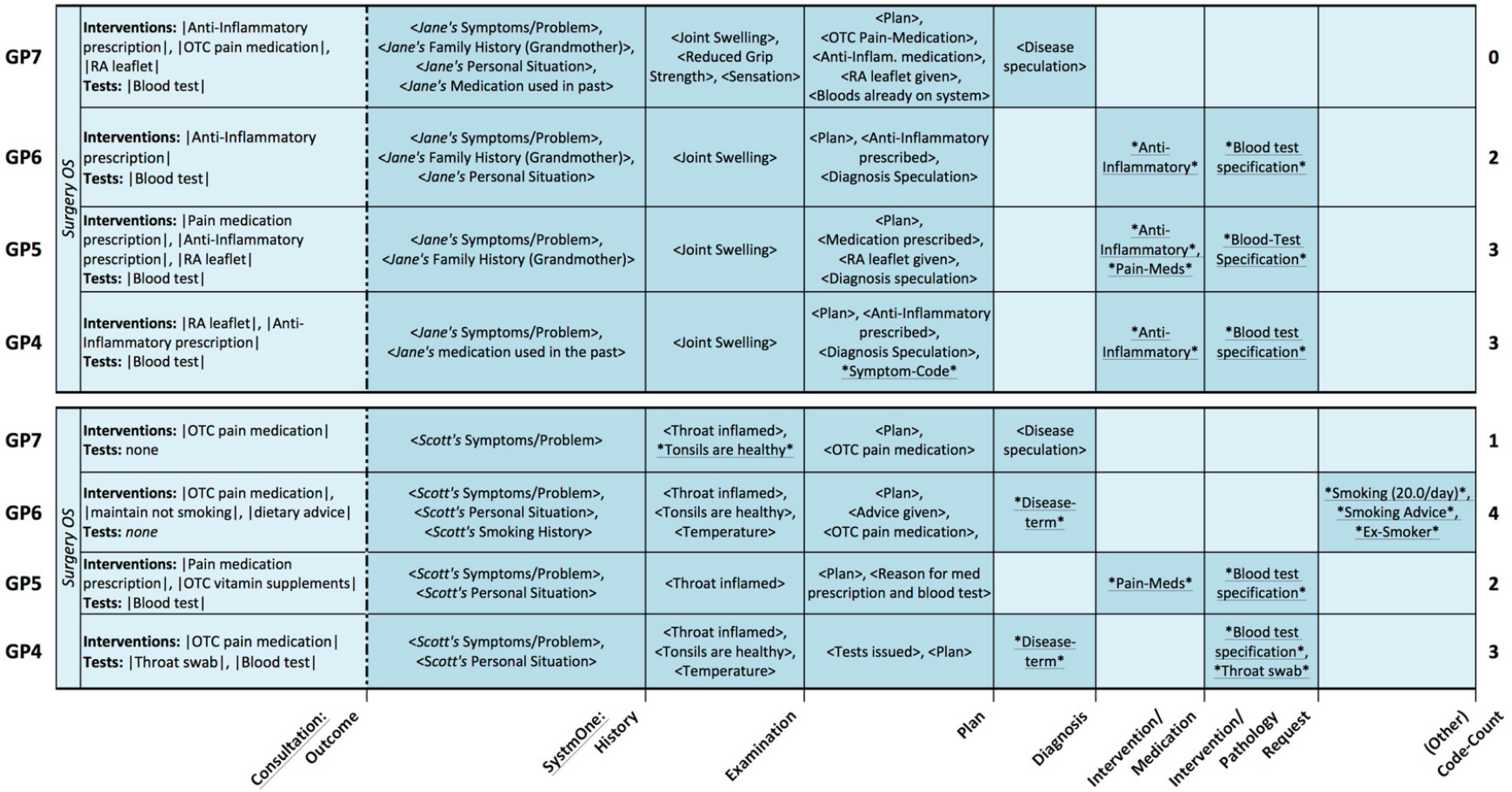


Figure 5.15: Content landscape of SystmOne EHR documentation of *Scott*'s (bottom) and *Jane*'s case (top). <>.. Free-Text, **.. Code, ||.. mentioned in consultation.

Coding

Medication, blood tests and terms denoting a disease would be coded consistently, while further details concerning patient background or examination were hardly coded. GP7 coded the least, documenting her diagnoses as free text only, using one code with *Scott* and none with *Jane* 5.15.

As described in the previous section, SystmOne continuously recommends codes to the users, while they enter the documentation. This exposed GPs to a high-frequent stream of codes they would pay little attention to:
GP6, in a feasibly QOF-led motivation, wanted to code the fact, that *Scott* was an ex-smoker and made use of SystmOne's code-recommending context menus (5.6) that had picked up on her typing "Ex-smoker" in the *History*-category. Thad yielded a template (fig?) GP6 had to navigate through, finally leading to three codes: 1) "*Smoking (20.0/day)"*, 2) "*Smoking cessation advice given*" and 3) *"Ex-smoker"* (5.15). This information does not represent the facts, as *Scott* 1) is an ex-smoker and 2) GP6 therefore didn't advise him to stop. Coding, in this case, has produced three inaccurate and contradictory codes. While 1) most likely is the result of a confusing user interface and interaction systematic, 2) might also have been issued knowingly in the QOF-context surrounding smoking habits by patients.
Another questionable code was issued by GP7: When filling out the *Examination* category on the throat examination, SystmOne recommended a tonsil-code she used without paying much attention to it: "*O/E tonsils normal/healthy*". It stands to reason, that this code is not in line with her intention at the time, since she would use the *Diagnosis* category to document her suspicion of a tonsil inflammation (*"? viral tonsillitis"*).

Tests and Medication

In both scenarios, all GPs who would deem it necessary to task a blood test (except GP7, who, in *Jane's* case, refrained from it, because she saw a blood-test was already set up in the system for her), would do so by using the *Intervention/Pathology Request*-function, which would code the blood test in the process.
All GPs who would decide to prescribe medication would do so by using the *Intervention/Medication* function. Except GP4 (ST), the GPs would generally document their recommendations for over-the-counter medication as free-text (ST: GP5, GP6, GP7), either in the case of unspecific (GP5/ST, recommending *Scott* to take vitamin supplements) and specific recommendations (GP7/ST, recommending *Scott* a specific type of over-the-counter pain-relief spray).

Word Count

GPs used the *History* category the most, followed by *Plan* and *Examination* (see fig. 5.16). Word count averages show GPs using slightly more space to describe *Jane's* case (*Scott*: 90,75 words, *Jane:* 103,75 words).

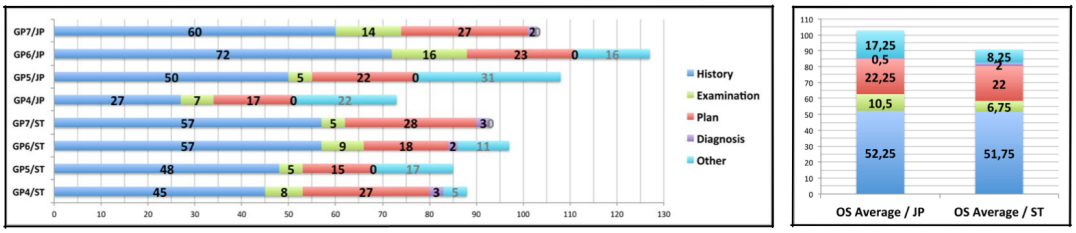


Figure 5.16: Word count of free text categories in SystmOne/surgery OS.

5.2.3 Vision

As with EMIS LV and SystmOne, the respective figure 5.17 yields an overview of category usage with Vision.

Category Usage and Coding

As described in the section on Vision's system overview, the categories used to structure the EHR documentation in Vision are directly related to a corresponding subset of the Read Code list, with the choice of a Read Code as a starting point of an EHR category entry.
GPs, as illustrated in figure 5.17, mainly used the **History** categories for documenting the consultations, which were offered to them on the basis of choosing a symptom-denoting Read Code, e.g. "Sore throat symptom". While possible to further disambiguate the category type as e.g. "History / Symptoms" or "History / Diagnosis", most GPs utilized "**History / Symptoms**" as category for describing *Scott's* and *Jane's* personal situation, symptoms and medical problem in free-text. GP11 also used it to mention *Scott's* past as a smoker in such way, GP13 also mentioned *Jane's* family history of arthritis there. Notably, a consultation record with Vision can contain more than one entry of a certain category, as evident with GP13, who documented two "History / Symptom"-entries: One attached to the Read Code "Stress at work", describing *Scott's* environmental stress factors, and one attached to "Sore throat symptom", describing his symptoms/medical problem at hand.

The category **History / Diagnosis** was also used by most GPs, but with more variability: GP12 would, with both scenarios, attach it to the organizational Read Code "Clinical management plan agreed" and utilize it to document the further plan of action agreed upon with the patients in the consultations. GP13 would attach a disease-denoting Read Code and use it for speculation on the underlying disease as well as to list the tests prescribed and advice given to the patients in the consultations. GP11 would use it as an all-purpose free-text field, attached to a symptom-denoting Read Code (fig. 5.17).

The category **History / Examination** was also used by most GPs and with little variability. Not unlike the *Examination* categories of SystmOne and EMIS LV, GPs

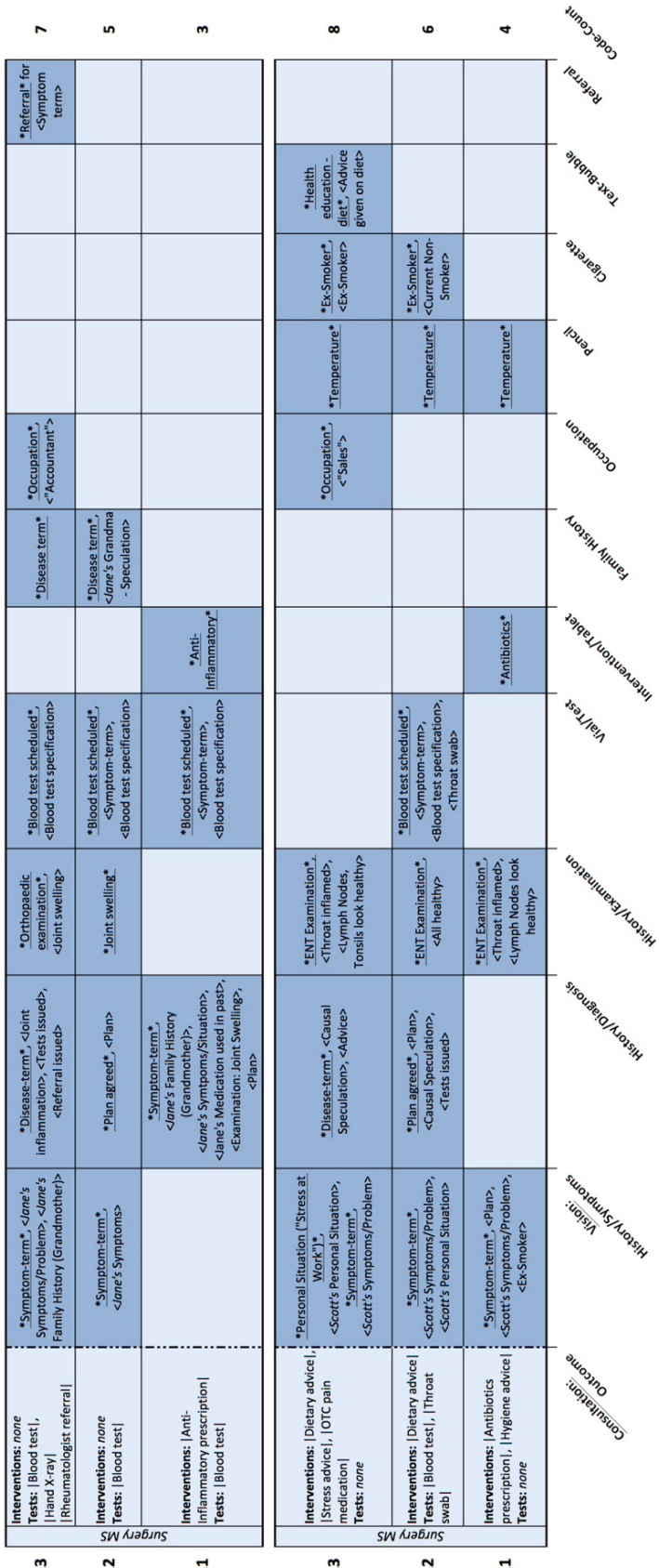


Figure 5.17: Content landscape of SystmOne EHR documentation of *Scott*'s (bottom) and *Jane*'s case (top). <>.. Free-Text, **.. Code, ||.. mentioned in consultation.

would utilize it to document their examination results, attaching them to an organizational Read Code, e.g. "ENT Examination".

All GPs of surgery MS documented *Jane's* family history of arthritis. GP11 did so with free-text in his all-purpose *History / Diagnosis* category, while GPs 12 and 13 issued a specific family history Read-Code, causing the usage of the **Family History** category.

As with EMIS and SystmOne, GPs of surgery MS would also document *Scott's* smoking status. GP11 mentioned it as free-text in the *History / Symptoms* category, while GPs 12 and 13 used the Read-Code "Ex-Smoker", causing the usage of the "**Smoking Status**" category (denoted by a Cigarette icon). GP13 furthermore would also code the occupation of the patients, using the Read-Code "**Occupation**", causing the usage of a category of the same name.

Further category usage is discussed in the following sections:

Tests and Medication

With Vision, GPs documented the blood tests issued by them consistently within the "**Test**" category, attaching it to the procedure-denoting Read-Code "Blood test scheduled" and further specifying the test in free-text. Notably, GP11 and GP12 would add a symptom-term in free text (see e.g. figure 5.11) to the blood test related free-text.

Notably, all GPs of surgery MS would test *Scott's* body temperature as part of the examination and document the result (*Scott* would say it was normal) as a test result, by using an appropriate Read Code, causing Vision to suggest the **Test Results** category.

As with other recommendations and advice issued by the GPs of the other surgeries, the GPs of surgery MS would inconsistently document them. GP13, who gave dietary advice as well as stress-related advice and recommended the usage of over-the-counter pain medication to *Scott*, documented the stress-related advice as free-text in the *History / Diagnosis* category, Read-Coded the dietary advice attached to a Read-Code of "Health education - diet", but left the OTC pain medication unmentioned. The throat swab test, which, with the GPs of surgery MS, was issued only by GP12, was documented as free-text as part of the blood test specification.

GP13 also issued a rheumatologist referral for *Jane*, doing so by the *Referral* category of Vision. He also issued a hand X-ray, but mentioned it only as free-text in the *History / Diagnosis* category.

Word Count

Figure 5.18 illustrates category usage by the GPs of surgery MS by word count. Notably, *Scott's* case caused the GPs to dedicate significantly more words to the *History / Symptoms* category, while, with *Jane*, *Symptom / Diagnosis* would be given more space. This might hint at a general understanding attributed to these categories by GPs 11-13. While potentially interchangeable, the *History / Symptoms* might be considered more appropriate to give space to for cases without a suspected diagnosis in the background,

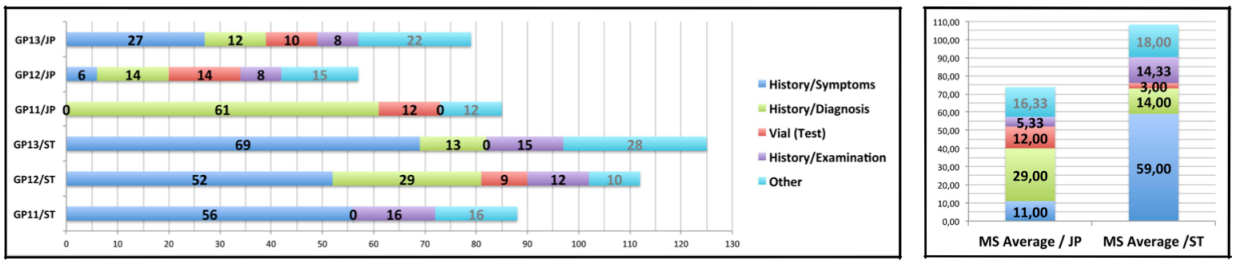


Figure 5.18: Word count of free text categories in Vision/surgery MS.

while *History / Diagnosis* might be considered more appropriate for documenting cases with such suspicions indeed existing.
Notably with GPs from surgery MS, *Scott's* case would trigger EHR documentation of longer average word count than *Jane's* case, while with the other surgeries, the scenarios would exhibit little variability in average word count across scenarios.

5.2.4 Comparison

Comparing *Jane* and *Scott*

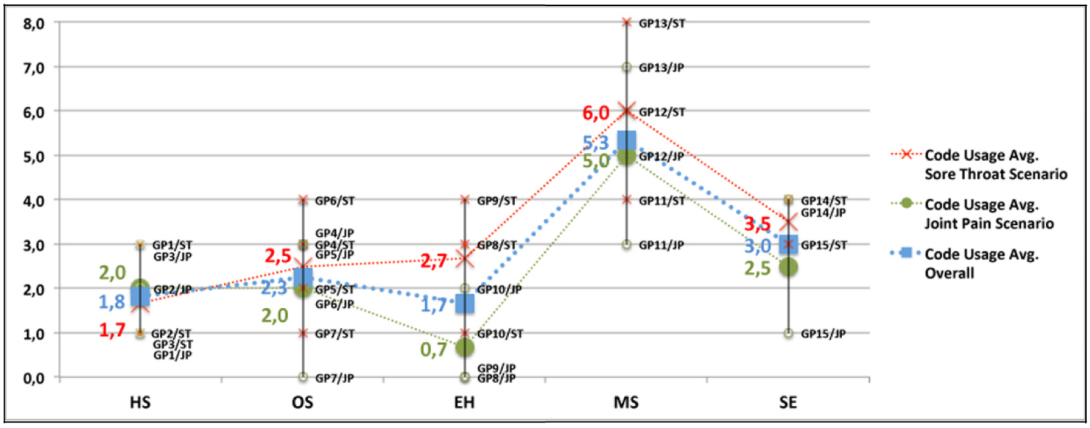


Figure 5.19: Overview of amount of codes used in EHR documentation across surgeries. EMIS LV (Surgeries HS, EH and SE), SystmOne (Surgery OS) and Vision (Surgery MS).

Notably, across all three EHR systems and four of five surgeries, *Jane's* case would be coded consistently fewer than *Scott's* case (see fig. 5.19) and the documentation would be divided in fewer categories (see fig. 5.20). Furthermore, slightly fewer words would be used to describe her case in three of five surgeries (see fig. 5.21).
In the walkthrough interviews, GPs hinted on a possible explanation in how they described *Jane*'s case and their medical plan. While all GPs described suspicions of arthritis during consultation and/or in the walkthrough interviews, two circumstances caused them to

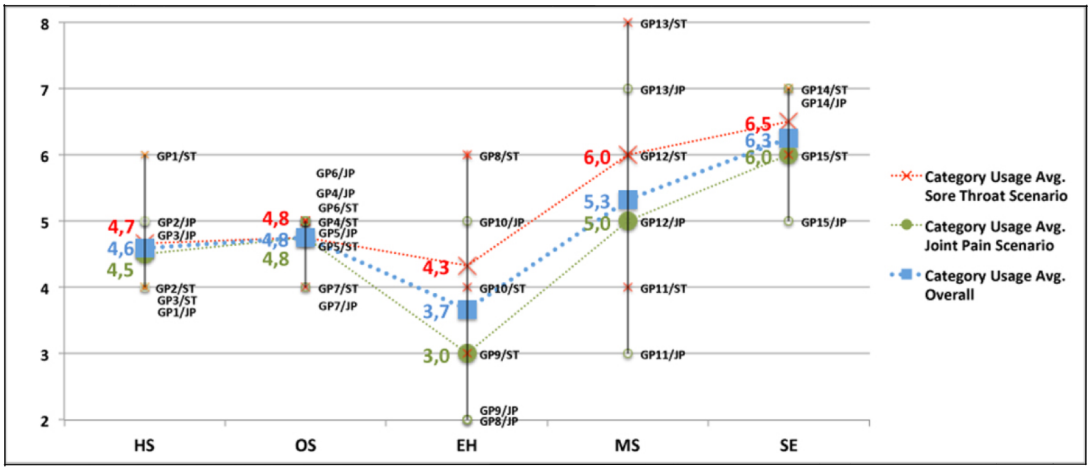


Figure 5.20: Overview of categorization of EHR documentation across surgeries. EMIS LV (Surgeries HS, EH and SE), SystmOne (Surgery OS) and Vision (Surgery MS).

keep the EHR documentation short of content: 1) Given the ramifications of diagnosing a patient with an auto-immune disease, they avoided being overly outspoken about their suspicions (see also: chapter 6.1.3 on codes as statements). GP1 didn't mention it, *"because it is her first episode and I didn't have a diagnosis at this point"* (GP1). 2) With her case, it was clear that she would consult again shortly with blood test results and only this evidence would warrant further steps. GP9, for example, *"didn't code anything because it's completely unclear. The blood test will show"* (GP9). GP4 would keep the record short, *"because this is a work-in-progress with her"* (GP4). GP12 would explain it similarly: *"[I put in] very little.. because. don't know.. I think it was enough because I'm seeing her again next week. That's enough information to tell me she's got rheumatoid, and I've to see and confirm it. .. a diagnosis will be put in with results back"* (GP12). GP9, apparently generally leaning towards more extensive and detailed documentation (see word-count overview, EMIS section), took an opposite position: *"This case certainly might end up with a referral in a fairly short order.. so I'm putting probably quite a lot on there straight away.. if I thought it was entirely self-limiting, then it might be much more basic"* (GP9).

With *Scott*, some GPs would describe the amount of EHR documentation as above average. This theme emerged with two explanations:
1) Several GPs couldn't quite make sense of *Scott* and were not sure whether he was being honest and upfront about his medical problem. *"Scotts history.. well I was wondering whether he was coming to me for his respertory infection or whether he was coming to me because he was stressed, had depression or something else going on. to remind me that when he comes back, I've had a bit of a suspicion about why he was coming "* (GP1). *"I've put in more than I normally would for a sore throat, because he is a bit more complicated, .. he's worried about actually the frequency.. although each individual episode is not*

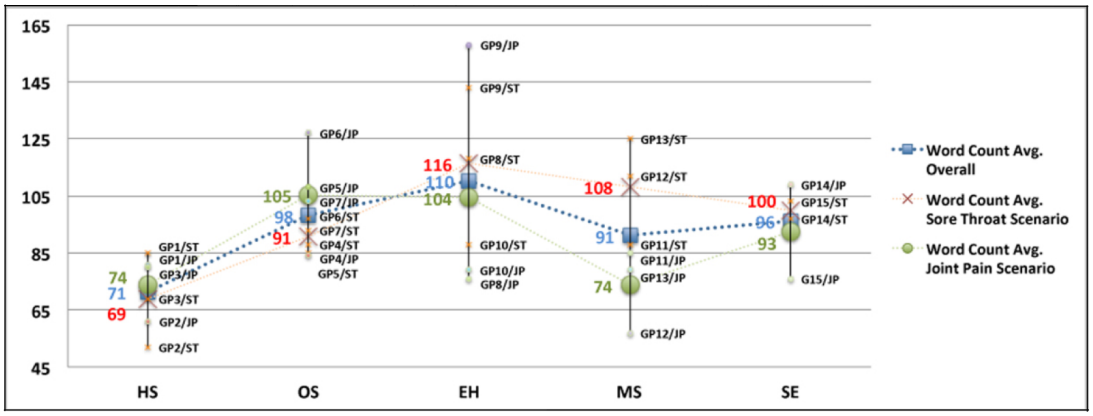


Figure 5.21: Overview of EHR documentation word count across surgeries. EMIS LV (Surgeries HS, EH and SE), SystmOne (Surgery OS) and Vision (Surgery MS).

severe" (GP3). *"Quite often.. there's some hidden agenda with it.. usually those sorts of things there's hidden agenda and they're scared if they have throat cancer"* (GP15).
2) Some GPs would situate his problem more in a focus on his environmental stress situation, invoking a more verbose practice of documentation. *"If it's a complex social thing, sometimes you need to put in more.. or if it was somebody coming in with depression they had loads of stuff that all were exacerbating their symptoms and I sort of.. going through divorce, this, that, cat died, .. and I'd write it all down.. next person thought.. ok yeah.. I know what I've got to deal with"* (GP7). GP5 also compared *Scott's* case to psychiatric cases, which usually are *"quite a long read"*.

5.2.4.1 Structural Differences Across EHR Systems

In relation to the different approaches to structuring a medical record and including Read codes taken in EMIS LV, SystmOne and Vision, documentations of *Jane*'s and *Scott*'s cases took different shapes: Documentations put down in Vision contained more Read codes than ones done with EMIS LV or SystmOne (fig. 5.19). Read codes with Vision were more diverse: While Read codes used with EMIS LV and SystmOne were mostly disease- and symptom-related, records created with Vision contained also procedure-denoting Read codes (e.g. "Clinical management plan agreed" or "ENT Examination") and more generally more densely represented details of the two scenarios at hand in codes, e.g. *Scott*'s body temperature.

In line with the system design of SystmOne, which, unlike EMIS LV and SystmOne, has a defined set of suggested categories to be used, documentations put down in SystmOne exhibited least variability in the amount of categories used across GPs (fig. 5.20).

Generally, the content analysis of EHR entries made with SystmOne, EMIS LV and Vision exhibited structural differences in line with the design peculiarities inherent to each system (see section 5.1).

5.2.4.2 Differences in Documentation Practice Across Surgeries

The content analysis furthermore yielded variability across surgeries. This perspective was possible with EMIS LV only, as only with this EHR system and the surgeries HS, EH and SE, a comparison was possible.

Figures 5.20, 5.19 and 5.21 suggest that surgeries may have internal norms on how 'verbose' GPs should be when documenting a case or which categories should be used in the process, hinted on e.g. by the comparably low variability among GPs of surgery HS in word count. This was described by several GPs; it is discussed in more detail in chapter 6.3 on different notions of what consists a 'good' Electronic Health Record.
The existence of per-surgery differences in usage practice of EHR systems was also hinted on by the category usage with EMIS LV (fig. 5.13); revealing different strategies in documenting different aspects of *Jane*'s and *Scott*'s consultation details, e.g. surgery SE used more and different categories than surgery EH.

5.3 Summary and Conclusion

This chapter described the three different EHR systems used by surgeries HS, EH, SE, MS and OS in how they structure information and involve Read codes in the process of documenting consultations. Furthermore, the results of a content analysis of the medical records resulting from the consultations with *Jane* and *Scott* are described, compared and interpreted.

Key Take-Aways

1) In their approach to assisting GP work, **EMIS LV, SystmOne and Vision structure and display information and integrate Read codes differently**. This encompasses different notions of a) What categories are available to the GP, b) How GPs can enter Read codes and c) How free text and Read codes relate to one another.

2) EMIS LV, SystmOne and Vision, according to GPs, come with **different advantages and disadvantages in usage practice** and they feature **peculiarities** that entail a necessity for different skills in GPs. For example: Vision's code-centric design requires GPs to know a specific Read code vocabulary. SystmOne's code prediction feature is challenging for GPs who are bad at touch typing.

3) Content analysis of the EHR documentation from the consultations with *Jane* and *Scott* revealed that the differences across EHR systems resulted in system-related **differences in the resulting usage of Read codes and structuring of consultation documentation in general**.

4) A comparison across surgeries hinted on **different norms** on documenting consultations with EHR systems **across surgeries**. For example: GPs from surgery HS are less verbose in their reports than others. GPs from surgery EH use fewer categories in their reports than others. This is described in more detail in the following chapter 6.

Relation to the Next Chapter

Picking up from the notion of norms and habits inherent to a surgery in relation to working with EHR systems that was hinted on by the content analysis done in this chapter, the next chapter takes a closer look on how GPs perceive working with free text and Read codes as well as what influences their concept of keeping medical records with EHR systems in general.

CHAPTER 6

Working with EHR Systems

In addition to variability in the treatment of *Jane*'s and *Scott*'s scenarios that emerged from a video analysis of the consultations and is described in chapter 4, and variability that emerged from a EHR content analysis that is described in chapter 5, this chapter takes some distance from the simulated patients' scenarios and describes themes of practice with EHR systems in general as they emerged in the analysis of the walkthrough interviews. Section 6.1 (*On Coding: Codes Are Useful*) describes themes of usage practice and variability concerning Read codes. Section 6.2 (*On Free Text: An Important Narrative*) reports on such themes regarding the usage of Free Text. Section 6.3 (*On Keeping Electronic Health Records*) describes usage practices and variability emerging around the task of keeping electronic health records in general. Section 6.4 (*Summary and Conclusion*) summarizes this chapter and formulates key take-aways.

Introduction

The task of keeping electronic health records is a complex one that depends on experience and is influenced by intricacies within the GP profession in general. The themes present in this chapter were chosen either when it would stand to reason, that they could point towards variability in the resulting EHR documentation, or when they would shed a light on the inherent richness in EHR usage practice in general. They are meant to be pointers towards interesting aspects in GP practice with EHR systems, illustrating the complexity inherent to working with computer systems in primary care.

This chapter is structured along the two building blocks of electronic health records: codes and free text. A thematic distinction emerged in the GPs accounts of their opinions on and experience with them: Read codes are *useful* and free text is *important* to GPs. These themes are taken as point of departure into details of aspects of practice surrounding coding and free-texting. After that, a third section on *keeping electronic health records* attempts to depict further influencing factors on the GPs' work with EHR systems.

GP10, on the difference of codes and free text: *"Depends on what purpose.. free text is what we read, but I guess if were going back and looking through your all sore throats, then coding is quite important"* (GP10). *"Free text gives you the finer details you never get from the code.. can be vitally important.. coding is very useful, helps you look back more easily, see the thread"* (GP11).

6.1 On Coding: *Codes Are Useful*

6.1.1 The Quality and Outcomes Framework (QOF)

A major theme of influence on coding behavior in GPs is the Quality and Outcomes Framework (QOF).

> The QOF is the annual reward and incentive program detailing GP practice achievement results. It rewards practices for the provision of quality care and helps standardize improvement in the delivery of primary medical services.
>
> It is a voluntary process for all surgeries in England and was introduced as part of the GP contract in 2004. The indicators for the QOF change annually, with new measures and indicators been retired. For 201/15 (sic), the QOF awards practices achievement points for:
> * managing some of the most common chronic diseases, e.g. asthma, diabetes
> * managing major public health concerns, e.g. smoking, obesity
> * implementing preventative measures, e.g. regular blood pressure checks [1]

If GPs deem a certain type of information QOF-relevant, there would be a significant incentive to code it: *"What I code.. is stuff that we get audited on for QOF.. and stuff that you might want to audit clinically on"* (GP4). *"In QOF, [codes] do get linked and things.. we're getting aware that we need to code things.. more important for audit trails, payment, QOF"* (GP7). *"It is a bit dubious, I don't like the fact.. if we're told to code something.. and we're not 100% sure that's what it is.. but it's useful"* (GP7).

Finding QOF-relevant information constitutes a task many GPs described as having present in their minds during consultations: When asked on his coding behavior with *Jane*, G9 stated, that *"there didn't seem to be anything for particular templates"* (GP9). GP8, on the same question: *"I **managed** to get in.. that he hadn't smoked. It's a QOF thing"* (GP8).
Looking out for QOF opportunities is part of the informal training GPs receive by their peers: *"With QOF, the senior-partner comes and tells me if I've done something wrong.. always watching..."* (GP12). *"I'm not massively QOF-motivated, because I'm not a partner, but I do try to keep things up to date. I do try to stay away.. aware of the QOF."* (GP5)

[1] Health and Social Care Information Centre (HSCIC): http://www.hscic.gov.uk/qof - last accessed 15-January-2016.

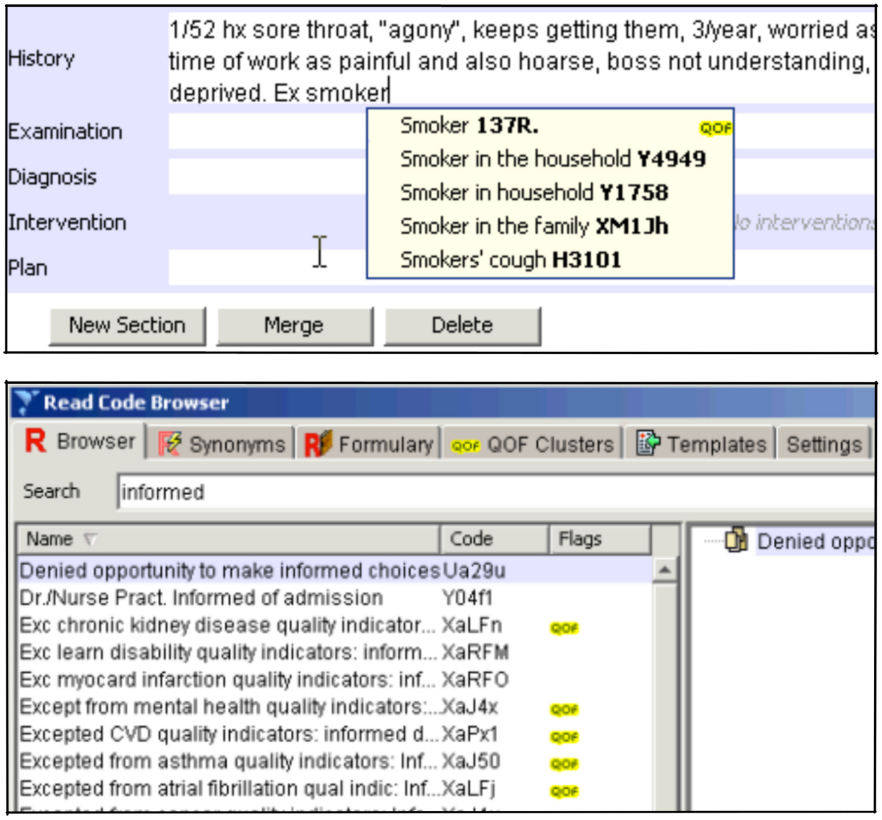


Figure 6.1: QOF flags. SystmOne supports highlightig QOF-relevant codes.

This mind-set, without doubt, affects GP practice and the subsequent EHR documentation. In a survey among GPs in the United Kingdom conducted in 2013, "97% of GPs reported bureaucracy and box ticking had increased since 2012 while nine out of ten GPs felt this took them away from spending time attending to patients needs. Eight out of ten reported target chasing had reduced routine available appointments to patients" [45, p. 2]. *"QOF-codes have implications for our ability to treat targets. And sometimes it's easy to mistakenly put a code in"* (GP4).

6.1.2 Codes as Organizing Elements

Several GPs described the usefulness of codes as organizing elements, mainly for two purposes: 1) As *Problem Titles* that would be attached to a patient's electronic health record, 2) as tokens for searches across the EHR data set of a patient, a GP or a surgery.

6.1.2.1 Codes as Problem Titles

"The purpose of a code.. Why I would read-code that.. So we could then use that as a problem to link to subsequent consultations. And retrieve it quickly" (GP4).

Problem titles allow to be linked and, by that, can structure a patient's medical history across consultations and along different medical reasons: *"The purpose of a problem is that it appeared.. in this list of problems here.. [and] subsequently can be found again. [...] It could be useful to code.. to make it a minor problem.. so you can link them. So, if you're making it.. current event type: minor.. it just keeps it on for a month"* (GP4). *"I always change it to a shorter time"* (GP10).

To GP4, Problem Titles are significant diagnoses: *"I would only make it a problem if it's useful to have it there.. every sore throat is not gonna go up there.. but for instance the next lady.. I haven't coded anything. But should our suspicions come to light.. she would definitely have her diagnosis coded, because it's more significant"* (GP4). The usage of Problem Titles is not obvious and develops over time: *"I'm still learning when to put it as a problem"* (GP7).

6.1.2.2 Coding Searchability

Coding *"structures the consultation.. it's quite useful to if you go back to actually look at it and find the information quite quickly.. it organizes it, structuring it"* (GP13).
Several GPs described coding as an act of bringing searchability into their EHR data: *"So I haven't Read-coded this one because it's a first presentation and it was too early.. I suppose I could have put a symptom.. yeah.. could have put Early Morning Stiffness.. it is not relevant at the moment, but you never know when you want to do a search or something.. on ..hand stiffness"* (GP8). *"You can't really search on free text..so as far as I can see the coding is to enable you to have power over your data.. erm.. not over your patients.. but.. so that we could generally say, you know, we have given smoking advice* (GP4).

"Patient for instance, would come in.. had a breast lump.. and looking back through the last couple of entries, I tried to reassure her that it was okay, and it wasn't really a lump at all, it was just, .. she had been a little paranoid. Just typing in 'breast' and coming up with all the consultations that she's been in, finding out that actually somebody else had reassured she has been investigated for it.. and it's all been found to be normal.. was all quite reassuring.. ultimately doesn't change my management in some ways.. but it's just really useful being able to pull up all the entries like that" (GP7).

6.1.3 Codes as Statements

Another theme attributed to the usage of Read codes by the GPs is one of *officialness*: codes are seen as statements of truth, verdicts and judgments that label the patient. With *Jane*, this led some GPs to be hesitant about putting their speculation into codes. As GP4 puts it: *"So I've put in arthralgia, which is a symptom code, as in painful joints, but in fact I was looking for polyarthralgia, something poly, saying several.. I haven't actually diagnosed rheumatoid, but it's one of my differentials. in the meantime I've put a symptom thing. because it has implications. If you have rheumatoid disease some people would put the code in and the problem this time.. but you'll find if you do an insurance*

report.. suddenly that code gets picked up and sent off .. and it has implications.. 'Oh, you've got rheumatoid!' So I use a symptom code here. [..] I'm not gonna put my neck out and say she's got RA probably until she's got a letter from a consultant saying that.. then it's coded and then it's a major problem." (GP4).

"I just don't feel confident in my ability to say "yes it's definitely that, let's code it".. it does have ramifications. Audit trails and things" (GP7). *"At the moment it's a symptom, we haven't got any sort of a diagnosis.. I'm not gonna label her as any particular type of arthritis or anything"* (GP11).

Such hesitation relates to a general reservation of GPs in relation to data sharing via EHR systems: *"I guess my concern is that notes are increasingly.. I mean the concept of confidentiality of computer notes is very open to doubt.. I find it difficult to reassure the patients.. when we [discuss sensible topics] how confident we can be about it. And that's a dilemma at the moment."* (GP2)

6.1.4 Finding the Right Codes

With the Read code database, the process of deciding on what code to use to express a medical issue at hand is often open to many possibilities and redundancies: *"The problem is, because the read code thesaurus is so huge, there is a lot of inconsistency in the codes that can be chosen. So you'll find that some people code sore throat, tonsillitis even when it is not really that. Consistency across the profession is very low"* (GP4).
This inconsistencies also reflect different approaches to the rationale behind coding. When explaining her coding of *Jane*'s scenario, GP14 explained: *"I put in arthralgia, and then I picked the one that sounded most like what she was saying"* (GP14). As GP3 put it, who used the Read code *"Sore Throat Symptom"* on *Scott*'s case: *"I could have coded that it's viral because that's what it was.. but that's really more what he presented with"* (GP3).
The choice of Read code thus can also be influenced by a motivation to find "speaking" Read codes that, in the opinion of the GP, might not be entirely accurate or contain all information, but overall encompass the patient and their problem.

Both with *Scott* and *Jane*, GPs tended to choose Read codes denoting symptoms rather than diagnoses. In relation to the "officialness" of Read codes, GPs would document speculation in free text; codes, as some described, need evidence. GP9 on coding his suspicions with *Jane*: *"If I'm referring someone with upper intestinal problems.. i know they're gonna have an endoscopy.. I would wait for the proof of the pudding, so to speak, and that would go on [the coded record] at that point"* (GP9). Same with GP2: *"I put a high-level code at first, then with test results coming in, [it would be] more specific"* (GP2).

In addition to the theme of documenting uncertainty by choosing symptom-denoting Read codes and place speculation in free text, as prevalent among most GPs with *Jane* and *Scott*, some GPs would express uncertainty by annotating Diagnoses with question marks. As e.g. done by GP7, a self-declared *"big fan of question marks"*; she documented her speculation as free text ("*?? arthritis*" and ("*? viral tonsillitis*").

6.1.4.1 Chance in Coding

Not least, the act of applying Read codes comes with a certain element of chance. Two examples, recounted by GP1 and GP4 on their consultations with *Scott*, illustrate mishaps in searching for Read codes that would affect what gets coded: *"I put bacterial, to start with, but it just came back with bacterial vaginitis *laughs*"*. Irritated by this result, she dropped her intention of putting a Read code denoting a bacterial infection: *"So I put in something elusive [Upper Repertory Infection NOS] and would recode it when test results come back. I suspected it to be bacterial. If viral, I would not investigate further"* (GP1). With GP4, a typographic error brought up suggestions that were not part of her initial intention, but fit anyway: *"I put 'viral sore th' and it came up with pharyngitis.. I thought 'That'll do!' [...] I spelled viral wrong ["vyral"].. had I not done that.. I might have ended up with a different code! *laughs*"* (GP4).

Finally, some GPs would describe a general element of arbitrariness and chance in relation to coding: *"You know with read-coding, you'll notice a massive difference.. in a way, for those, there's loads of things you could have Read-coded, and I'll usually just gonna go for whatever comes.. I think if you had caught me on a different day, I would do something differently. I'll put in different things on different days"* (GP1).

6.2 On Free Text: An *Important* Narrative

As described in the introduction of this chapter, a frequent notion among GPs when describing free text would be *importance*. *"Free text gives much more details and granularity to the flavor of the patient and the consultation and also the clinical interaction. The code.. just represents.. opportunities for data retrieval.. you can search those codes, you can organize information around those codes, but for the most part, they don't lend useful clinical information"* (GP4).

"Part of the duty of a doctor.. to keep good records. for various reasons.. it shows what process I've gone through, for partly medical-legal, partly if anyone else sees him about this, to show what my thinking is, which.. and so it's to record as accurately as I can, what he is telling me.. the problem is. including what impact it has on his life" (GP10).

6.2.1 A Narrative For Assisting Memory

A frequently described notion about free text is that it supports the GPs memory: *"I think the free text really adds the information around the code. you can't really tell a story just by read codes"* (GP4).
This story serves to describe a patient by more than just medically relevant information, painting a picture of a human being as a whole: Free text, according to GP2, serves to document *"history, remember a few odds and sods about them.. so.. you know.. the concept of illness vs. disease.. illness being the patient plus their context. instead of next time jumping in with 'Oh, you're the patient with the joints'. [I would] put in that he's an accountant.. bits that would remind me"* (GP2). *"[...] so I might put something in*

like son-in-law's wedding coming up.. which is not medically interesting.. but helps my memory" (GP4). Free text, also according to GP11, is for *"recall.. you put something in there that will identify that patient. Personal info. They come in.. how's the job going, has the wife, the dog died"* (GP11).

6.2.2 A Narrative For Assisting Reasoning

Another frequently described aspect of free text in EHR documentation is its capability to assist GPs with reasoning in between consultations. Explanations of decisions and plans of action contextualize further action and provide important information: *"I put in what he tried.. I needed all that information to justify what i did and my decision and, you know, should he come back to me, I'd like to look and see those sorts of things"* (GP4).
Similarly, information on the medical history also contextualize the medical problem at hand. GP10 describes his reason to document *Scott*'s symptom history: *"[It was the] first time he has consulted, it does have some relevance. So to say 'You did have those other two episodes.. and didn't consult for, didn't have any treatment for, they got better!'"* (GP10).

Setting up the patients story as a narrative in the EHR system and by that recounting the consultation and the medical problem at hand also helps with reasoning, as an example made by GP4 illustrates: *"Sometimes when I'm writing up something.. I've more or less described a DVT [Deep Vein Thrombosis] and I think.. oh my god I might get that patient back.. or even pick them up on the phone and say.. just one last thing"* (GP4).

"I tend to write down the duration of symptoms ,.. what symptoms were.. what was worrying them most.. ultimately it's a prompt for me.. when they come back.. as to why i was worried about them or what would be going on" (GP7).

"Things like patients symptoms, any specific concerns they had, usually put in something about the plan, if there's anything specific about the examination you can't code, I'll put that in, for example with children, I'll put in 'run in the room smiling', or 'playing happily'. [Interviewer:] Why? [GP11:] It's not the sort of thing you can code, but it's very important, it's the sign that the child is well" (GP11).

6.2.2.1 Negatives

A frequent occurrence within free text in the EHR documentation about *Scott*'s and *Jane*'s consultations were "negatives" - notes on symptoms or other aspects that were not present. GPs described those as important points for reasoning and characterizing a medical problem.
GP3, for example, when documenting *Scott*'s case, wrote that he is *"not overly prone to other infections". "I think i wanted to convey that he wouldn't have a known immune deficiency"* (GP3). Such type of information occurred frequently in free text concerning examination or anamnesis results for both *Scott* and *Jane*.

GP8 explains negatives in her free text documentation of *Scott*'s case similarly: *"That's just to show that I've considered depression.. and I've sort-of gone part of the way towards asking all the questions and then sort of dismissed that.. I've gone down that route and dismissed it.. it's important for the next person coming in that I've considered and didn't feel that was the diagnosis at the time"* (GP8)

6.2.3 A Narrative For Medical-Legal Purposes

A point closely related to documenting negatives for purposes of reasoning is the usage of free text for medical-legal purposes: It serves for *"recording what there wasn't.. „for defensive purposes"* (GP1).
"The free text, as far as I'm concerned.. is what you have to stand up in a court of law and read out.. so to me.. I have to be able to look and.. can remember the patient from it .. I have to have ruled out significant.. you know.. significant problems.. so I put down significant enough negatives" (GP4). *"Really medical-legal reasons more than anything else"* (GP11). *"Defensive purposes. if someone looks back and asks questions, in free text you can find a lot of information, really, why the decision was made. it explains the code"* (GP13).

6.3 On Keeping Electronic Health Records

6.3.1 Integrating EHR Into GP Work

The computer is a steady companion in general practice. In the cases of all 15 GPs, the placing of the offices' furniture and equipment, as well as the seating position of the GPs in relation to their surrounding, depicts a computer in a central position (fig. 6.2), hinting at the importance of this electronic device for GP work.

GPs mostly used the EHR system subsequently to concluding the consultation and saying good bye to *Scott*/*Jane*, to document the case. Some GPs would also use it intermittently during the consultation, for purposes of documenting the case, issuing prescriptions or printing out information leaflets on rheumatoid arthritis to *Jane*.

This intermittent use can disrupt the rapport between doctor and patient during a consultation, possibly causing the GP to miss details of the patient's account of the illness. From the point of view of the interviewed GPs, it nonetheless is necessary to relieve cognitive load: *"I don't write on the computer while they're here, unless there's loads of stuff"* (GP7). *"I find if I don't do it at the time, I tend to forget it. [..] But I didn't do any typing while he was talking.. I tend not to. Sometimes I do. After he'd left, I wrote up his history, and the examination findings"* (GP4).

The ability to accommodate both the need to document the consultation in fear of otherwise forgetting details on one hand and retain doctor-patient rapport on the other hand also varies along skills of working with computers: *"We're all different.. I think the registrars tend to put more.. because they.. are learning.. and they are quick.. at typing.. I'm quite a quick typer.. I've done a typing course.. that's helpful.. I do quite often type*

Figure 6.2: The computer as a handy companion in GP practice.

during consultations.. because my memory is not good.. I will just put some of it down... generally leave it to end (GP8).
GP14, who would apologize to both *Scott* and *Jane* during the consultations about her 'bad' typing skills: *"I am very slow at typing, and in fact frequently I say to patients, like I did to my two patients here, sorry about the typing, I'm a bit slow, and .. because it does .. that's one big problem that I would find.. is because I'm not a touch typist.. I'm not like there, as they're telling me.. I am .. they tell me, and then I've got to go back here and do it.. so my focus would come off the patient.. and a lot of patients would be sitting there, while I'm sitting there like that and they're gonna.. you know I don't think that's great.. definitely we all need touch typing skills. [Interviewer:] I don't think with touch typing you can completely... you know the computer takes time.. [GP14:] It does takes time.. and it's just.. everything they say you then have to get it on there.. and if you're slow.. it is time-consuming in a consultation.. it swallows a fair bit of your ten minutes.."* (GP14)

Time Pressure

Putting in EHR documentation takes up time and thus has to compete with seeing patients and other tasks of GPs, introducing an element of chance into the EHR documentation: *"I think it varies a lot as well, I mean if you're working under pressure in time, I will put in less or be far more cryptic"* (GP2). *"There is a risk [for something to be left off].. especially if you're behind your surgery.. and you have to hurry"* (GP7). *"The family*

history [(of Jane)].. I could've entered the family history. That could've been a Read code. In the speed of putting things in.. I didn't do that today" (GP8).

6.3.2 Experience and Literacy with EHR tools

As was hinted on by the content analysis of *Scott*'s and *Jane*'s EHR documentation (see chapter 5.2.4), surgeries develop their own understanding of what constitutes 'good' electronic health record keeping. This understanding is negotiated among surgery co-workers; it depends on the individual perspectives of the involved people, who exchange them and reflect upon their practice towards the construction of a collective understanding of EHR-keeping:

Having Individual Perspectives ..

In working with EHR tools, GPs bring their individual level of experience to the table: *"Influences on EHR usage? From a medical-legal aspect that's quite broadly promoted.. about the importance of record-keeping, .. erm.. and coming through.. a hospital.. that's very much what you.. I was involved in doing some audit-collecting for the clinical negligence scheme... going back a few years we were just looking at notes. I think there certainly general principles are picked up along the way"* (GP3). *"I've a habit of making it really, really long, because I'm still quite new.. I'm still working at the hospital where you have pages and pages.. and I'm trying to, sort of, cut down, what I actually write down"* (GP7).

.. And Exchanging Them ..

For themselves and also by sharing those experiences with fellow surgery members, knowledge and understanding shapes their work with EHR tools: *"Dr. [W].. in his past life.. he was involved in .. setting up codes.. he gave all the partners a sort of tutorial on.. history of Read codes.. why it's important.. what you shouldn't code.. then I saw 'Oh, I'm coding for a reason'"* (GP4).

GPs would describe that they pick up skills in working with EHR systems mostly 'on the job', meaning 'learning by doing' as well as learning from one another: *"I'm a GP trainer, and have been for quite some years.. setting an example to my trainees is a part of.. generally the ethos of having quality record keeping"* (GP10).

By several GPs, this is described as an on-going process, e.g.: *"[I had] training in my introduction-period, couple of hours dedicated sit-down. Did that with my fellow GPs. We still do that on a weekly basis. sort of on-going. There are still things coming up and I think: 'I didn't know I could do that!'"* (GP3). This does not necessarily mean, that GPs of one surgery have to have the same routines: GP1, who stated having been with the surgery for three years, on ordering a blood test with the Pathplus-addon in EMIS LV: *"Nobody has shown me how to do it yet.. so I'm still using the form"* (GP1). Obviously, GP1 manages fine without using Pathplus.

.. Towards a Collective Understanding of Electronic-Record-Keeeping

Some GPs described how 'good' use of the EHR system is also discussed among members of the surgery: *"Everyone should stick to the same codes.. we have practice agreement, we sometimes talk about codes and how we want things put on the screen, that's developing all the time, really"* (GP11, surgery MS). *"We have decided that.. arbitrary use of Read code is probably less helpful than actually thoughtful, considered use... So when there's a benefit to use that, we use it to aggregate information, than putting in a code for the sake of it"* (GP4, surgery OS).
"Sometimes, there is two read-codes for the same thing,.. that can lead to confusion.. we try to use when someone is hypertensive, we try to use hypertension, but then there is hypertensive disease, various other ones. We try and use the same ones.. but there's often too many different codes for the same thing" (GP11). *"Coding activities vary tremendously, you'll find someone trying to code everything 'cause they feel that's the purest way of expressing clinical information in a record. Others just free text everything."* (GP4)

Importantly, such collective understandings of good EHR-keeping affect which elements are more or less likely to be coded; a consideration influenced by how useful such coding is for the GPs: *"I'm not putting a read-code for sore throat, if the blood test shows he's neutropenic, that's gonna on there, cause that's immediately pertinent to a colleague, something they might have to see fairly quickly. [..] the 'Recurrent Sore Throat' is too vague to be useful, it's almost irrelevant you know, for colleagues* (GP9).

Age and EHR Usage

Several GPs mentioned having the impression, that older GPs would be less verbose in their EHR documentation. *"With time, notes get condensed. With confidence and experience, you learn to condense things. If you're somewhere only temporarily, it's much more important that your notes are very coherent for people taking over from you"* (GP3). *"Older GPs put less.. newly qualified put in more"* (GP12). *"Some people put generally less. They just put code, you have to guess what happened.. as older you get you put less text in it."* (GP13).

In the content analysis of *Jane*'s and *Scott*'s EHR documentation, this would not be visible as a significant trend, though (fig. 6.3).

6.3.3 Collaborative Work

When asked about how the EHR systems were used for purposes of collaboration, most GPs described scenarios of organizational tasks as well as collaboration with other doctors or nurses. For example: *"Sharing.. with all the other doctors.. so.. occasionally the admin stuff would look in the notes.. if they've had a prescription request.. and it says.. Dr. M reduced these.. but she hasn't put it on a prescription.. they might look and see.. oh, she did say that"* (GP4). *"The nurse will check when his blood-tests return. so the nurse would look at the record"* (GP5).

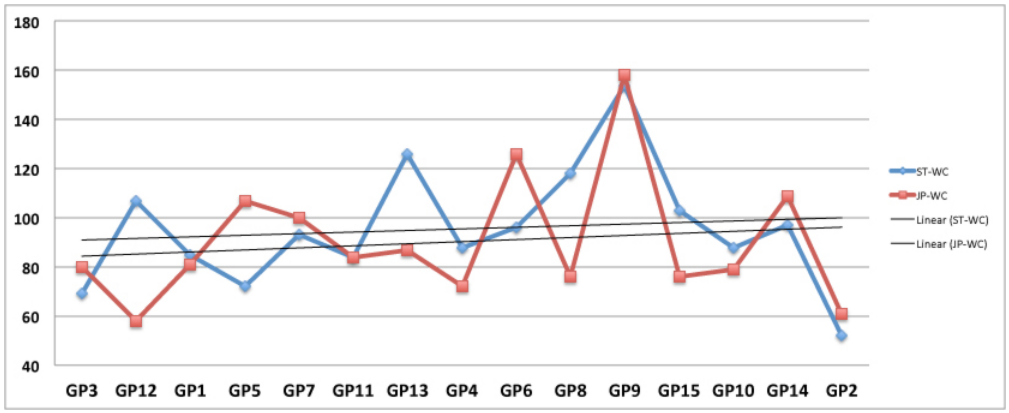


Figure 6.3: Word count with Jane (JP-Scenario) and Scott (ST-Scenario). GPs in order of age. An overview of the GPs with gender and age range can also be found in figure 3.2.

Routine tasks like ordering a blood-test are handled differently from surgery to surgery, depending on the individual organizational structure and habits of a surgery. According to the GPs descriptions, tasks would often also be handled distributedly. GP8, for example explained why he would choose to specify a blood test order in free text only: *"I've put in these blood-tests.. erm.. which I put on the computer because.. when [Jane] goes to have the blood-test done, then immediately Sue can see what needs doing. and its clear.. and that.. erm.. saves me having to do a form"* (GP8). *"I would normally use the Dictaphone.. and it will be done by the secretarial staff, they would do all the coding in relation to that aspect"* (GP14).

A Narrative For Collaborative Work

For collaborative work with EHR systems, both free text and codes play their own role. Again following the themes of usability and importance attributed to codes and free text by GPs, free text takes in an important role for collaborative work. It is used for communication and patient-related knowledge exchange among a surgery's GPs in front of the background that, on several consultations, a patient might not be attended to by the same GP: *"I think, these days, with more of us working part-time, she's likely to see somebody else.. it would be helpful for the next person to be able to quickly just go over that and then carry on rather than starting afresh"* (GP2). *"[With] clinicians, lot of people are part time.. we often see a patient after they've seen someone else for the same problem, and so the note-keeping is very important"* (GP11). *"I think it's good for a record for the next person so what they are going to do next and what's been done already, so, it's fine if he comes back to me, you normally have the triggers in your mind to remember, but it could be somebody else, so it's sharing of information really"* (GP4). Similar remarks were made by most GPs.

GP7 furthermore described, how the EHR system would also be used for her GP training:

"My Trainer would look at them, what my management plan would be, make sure I'm being safe.. and I haven't missed anything.. and I tend to sit down with her at the end of the day and go: right okay, we've seen these patients, attune my documentation and see if there's anything I need to improve on. from a training point of view, it's really useful" (GP7). This gives EHR documentation an important role, as it gets used to assess a GPs learning progress.

6.3.4 Retrospective Modification of EHR Documentation

This distribution of tasks would also cause the GPs to impose a certain usage practice not necessarily common or intuitive to others. GP10, for example, in terms of EHR usage practice, viewed himself as *"Untypical for the rest of practice.. I'm very much out on a limb.. I come across things my partners have written.. I try to be systematic.. put examination in examination and so forth.. should be easy to read and understand.. some of my partners don't do that"* (GP10).

In between *Scott* and *Jane*, GP10 received a task to do on the EHR system, relating to a consultation that was done (and documented) by another GP: *"While I did that.. I tidied up the note.. I mean you have stuff in there.. that shouldn't be there.. so that people have important diagnoses.. [so that you can] trust the front page more.. and you can use it! If I've got my notes neat, if I have a reasonable record, we've designed a standard letter, to make sense, so all I've got to do is send a letter to my secretary saying send letter to so-and-so. in the bottom I would say 'refer to cardiology for advice'. and .. I just send a patient note to Wendy: 'Standard letter C'"* (GP10). GP10 retrospectively changed an EHR entry so that it would be up to the standards of 'Standard letter C', possibly affecting meaning and content as intended by the attending physician of the respective patient.

GP2 described a similar behavior: *"With most patients I would go into the problem page.. and look at all the problems there.. and quite often.. this is my OCD-aspect, I'll tidy them up a bit"* (GP2).

GP14 recounted a situation in which QOF-related incentives motivated a larger, retrospective change of EHR documentation: *"We don't always all use the same codes.. well recently for example, there were several different ones for hypertension.. there's essential hypertension, there's hypertensive disease, there are several of them. and they all looked perfectly feasible, you know, but when you come to collecting data from the computer, and you have to do it for QOF purposes or whatever, it would not identify some of the codes we were using, so we have had to batch-transfer everybody with hypertension over to hypertensive disease so we were all using the same one.. but with so many thousands of millions of diagnoses, it is very difficult"* (GP14).

6.3.5 Influencing Factors When Facing Patients

In their work with *Scott* and *Jane*, some GPs would use the EHR system also to document an agreed-upon 'final' account of the consultation and the plan of further action. This

would bring up an interpretation of the computer as a sort of authority externalized to the consultation, that receives and documents decisions and 'makes them official'; decisions that would lend themselves to improve patient-doctor rapport. This relates to GP training e.g. as described in the Calgary-Cambridge guide mentioned in chapter 4 as step towards "shared decision making" towards improving subsequent adherence to the decision by the patients. [pt. 48][43]: *"I'm trying to.. have it mutually satisfactory.. so that the patient's agenda.. whether I would agree with it or not.. is somehow included"* (GP9).

When creating an EHR entry, GPs are mindful of the patient's view on EHR documentation, weighing the necessity to document certain information against possible ramifications in their relationship with the patients. As GP8 puts it: *"[We had a] surgery discussion.. what we would type in there, because that's what would get relayed to the patient. That sort of discussion would influence what I do.. It's difficult, because you can't rely on the patients either that they phone-up results, or remember that you told them to have a low-fat diet"* (GP8). GP2 and GP4 also described such thoughts: *"[..] relating to other possibilities for a sore throat, which I didn't discuss with him, which are of a more sensitive nature.. sometimes younger men with sore throat can have STDs"* (GP4). *„I think it's visible.. I trans-check it with them as I go along.. if it is slightly controversial, psychological or something.. I think it's quite useful to negotiate with them.. you know what their agreement.. if possible agree.. and if they don't agree.. agree that we can put on both our comments. .. that they are not an alcoholic or whatever"* (GP2).

6.4 Summary and Conclusion

This chapter presented themes on usage practice with EHR systems that emerged in the walkthrough interviews done with the GPs.

Key Take-Aways

1) The usage of Read codes is incentivized by several aspects of **usefulness** that are not directly medically relevant to the GPs' work, causing **differences in the likelihood certain information might be coded** with Read codes: (a) Information especially valuable for (a) QOF purposes or (b) structuring a medical record might be more likely to be coded.

2) Read codes might have the character of a label for a patient, causing GPs not to use codes for information that is considered very relevant, but also uncertain, as they find it compromises doctor patient confidentiality and might have unwanted ramifications.

3) The Read code database is extensive, containing lots of subtle differentiations that are considered unnecessary redundancies from the perspective of the work a GP does. Finding the 'right' Read code can be **influenced by chance**.

4) Free text is used to document **important** information that assists with memory, with reasoning, that acts as security for medical-legal purposes and supports cooperative

work.

5) In GP practice, **EHR systems are** used in a field of tension between the necessity to document information and do work on one hand and the time this takes away either from doctor-patient rapport during consultations and from GP work in general, **exerting time pressure**.

6) When working with EHR systems, GPs bring their experiences and viewpoints to the table. They **continuously develop and change their EHR work practice** alongside experience and the exchange among surgery co-workers. In certain aspects, they **align towards common surgery-understandings of 'good' EHR-keeping**.

7) GPs reported, that, under certain circumstances, **EHRs can be retrospectively changed** to adhere to requirements of automation or because they look messy. GP14 furthermore recounted a situation of retrospective change of EHR data en masse for QOF reasons.

8) When creating an EHR entry, **GPs are mindful of the patient's view** on EHR documentation, **weighing the necessity to document certain information against possible ramifications for their relationship with the patients.**

Further Discussion

The results of this chapter, along with the results of chapters 4 and 5, are discussed and related to one another in chapter 8.

CHAPTER 7

Secondary Analysis of Primary Care EHR Data

This chapter aims at relating the phenomena of practice and variability in work with EHR systems that emerged from the analysis of the PREP field study's data material to another area of interest related to computer-supported GP work: Secondary analysis of primary care EHR data.
To this end, the chapter is divided in two major parts: Part one establishes the approach of secondary analysts to primary care EHR data and consists of the following sections: Section 7.1 (*Definition*) characterizes the concept of secondary analysis, section 7.2 (*Data Sources for Secondary Analysis*) describes the data-sources for such analysis and section 7.3 (*Dimensions of Data Quality and Strategies of Validation*) describes notions of viability of and strategies of validation of such data. Part two, found in section 7.4 (*Phenomena of Variability*) discusses opinion, stance and awareness of phenomena of variability and their implications.

Information Sources

This chapter draws from related research literature; among others, several literature studies on aspects of secondary analysis with EHR data were taken into account: Herrett et al. [46], Khan et al. [9], Häyrinen et al. [47], Weiskopf and Weng [10] and Botsis et al. [12]. Furthermore, expert interviews were conducted with four academic epidemiologists and two experts in the fields of EHR design and Natural Language Processing (NLP). The report given in this chapter is based on the results of a thematic analysis given by the experts:

E1 Epidemiologist. Research with general practice EHR data.

E2 Epidemiologist. Research with clinical data on cancer incidence.

E3 Epidemiologist. Research for influenza surveillance.

E4 Epidemiologist and Health Services Researcher. Research with rheumatoid arthritis, also involved in design of the PREP field study.

E5 EHR researcher. Research on interoperability of EHR systems and applications for secondary analysis.

E6 NLP researcher. Research on Natural Language Processing (NLP) in EHR data analysis.

7.1 Definition and Applications

7.1.1 Definition of Secondary Analysis

Boslaugh gives a compact definition of secondary analysis as follows:

> If the data set in question was collected by the researcher (or a team of which the researcher is a part) for the specific purpose or analysis under consideration, it is primary data. If it was collected by someone else for some other purpose, it is secondary data. [8, p. 1]

Resource-economy poses a key advantage of secondary analysis, as the researcher does not have to devote time and other resources to data acquisition and collection. Secondary analysis furthermore holds an advantage in the possible breath of available data: While individual researchers will often not have the resources to conduct large-scale data collection regarding a particular interest, other institutions might collect rich data as a by-product [8].
Major disadvantages in working of using secondary data, according to Boslaugh, include (1) its situatedness in its primary collection incentive; since the collection concept of this data was not designed in line with the research interest of the researcher conducting its secondary analysis, relevant information may not have been collected or collected in a different way. Furthermore, (2) the distance of the secondary analyst to the original data collection and by that, a lack of knowledge and awareness of its practical particularities [8]. Boslaugh recommends certain questions for the researcher to ask themselves for the assessment of candidate data for secondary analysis:

> In particular, you will want to answer the following three questions:
> 1. What was the original purpose for which the data were collected?
> 2. What kind of data is it, and when and how were the data collected?
> 3. What cleaning and/or recoding procedures have been applied to the data?
> [8, p. 8]

In addition to such questions for reliability and validity of the data, also ethical concerns arise from using personalized and possibly sensible medical data for secondary analyses. As Heaton remarks, informed consent should not be presumed, and a professional judgment might have to be made on whether the secondary use goes beyond the initial consent of the involved people [48].

7.1.2 Applications with General Practice EHR Data

Secondary analysis of General Practice EHR data is a popular empirical instrument for various types of research, including public health research like epidemiology or Health Services Research (HSR), medication-related research or clinical trials[1]. Depending on the concrete research interest, secondary analysis can work hand in hand with primary analysis [8]. While the opportunities of and by that also interests in primary care medical data for secondary analysis are manifold, this chapter describes secondary analysis efforts along the fields of expertise of the interviewed experts - epidemiology and Health Services Research (HSR). Regarding primary care EHR data, they hold common requirements of validity and reliability that assumedly hold relevant implications for other secondary analysis scenarios as well.

Epidemiology constitutes research that "*deals with the incidence, distribution, and control of disease in a population*"[2]; investigating questions for disease incidence, prevalence, patient (co-)morbidity or mortality in pursuit of objectives as better understanding the nature and interdependence of diseases [49]. The interviewed epidemiologists described own projects of secondary analysis of General Practice EHR documentation for purposes of determining the comorbidity of asthma and allergies in children (E1), possible correlation of young parenthood and cardiovascular disease (E5) and influenza incidence (E3).

Health Services Research (HSR) denotes the "*investigation of the health needs of the community and the efficiency and effectiveness of the services provided to meet those needs*"[3]. As put by E4, expert on HSR: "*What you're often trying to describe is the relationship between health services and people with health conditions. Are they going through quickly? Someone with RA: Are they getting identified quickly at first? Are they getting the right treatments? Are they getting cured? Are they getting the outcomes that they want?*" (E4).
In this, HSR investigates "*international, national, and local health systems and their interconnectivities, and policies made and implemented at all levels of the health system*" [50]. For understanding the effects of health policies and systems, EHRs posed an interesting source of insight since long time: A 1990 EHR data study on preventive procedures in primary care argued for the usage of computers in general practice as it can be

[1] https://www.cprd.com/ObservationalData/, accessed 15-February-2016.
[2] http://www.merriam-webster.com/dictionary/epidemiology, accessed 1-Feb-2016.
[3] http://medical-dictionary.thefreedictionary.com/health+services+research, accessed 1-Feb-2016.

> "systematically and uniformly collected for analysis and feedback": "[It] means that the procedure takes almost no time [in comparison to the then common, *manual* audits]; the whole practice is covered without sampling; there is no difference in audit systems between practices or within a practice over time; and any block of patients revealed as lacking care - for example, children not fully immunized or women overdue for a cervical smear testing - can be identified with their names and addresses."
> [51]

E4 described the investigating prescribing-practice of antibiotics with patients of the Common Cold as an example for HSR, reflecting a tension field of the disincentivization of antibiotics prescriptions for patients with the cold by public policy on one hand and the nuanced, context-dependent question of whether patients "*get what they want*" on the other hand, as also described by Smith et al. [52]. Other exemplary HSR projects include the works of Dunn et al., who investigated antibiotics prescribing practice in relation to possible complications of tonsillitis [53], of Edwards et al., who investigated treatment-variation in rheumatoid arthritis [54], of Judge et al., who investigated the effect of guidelines on rheumatoid arthritis [55] or of Forster et al., who investigated the effectiveness of NHS health checks in the United Kingom [56].

7.2 Data Sources for Secondary Analysis

Enticing Richness of EHR Data

Turning to data originating from health care work with EHR systems for secondary analysis represents an opportunity to "*mine vast amounts of structured medical record data*" [57], it is enticing in several ways. Electronic Health Records represent a "full take" of a populations' medical data and by that open up a much more detailed view on a populations' health situation, also encompassing rare phenomena that are more elusive to classical epidemiological research, which otherwise operates with sample data and extrapolations on the overall population, as E5 explained. Furthermore, EHR data holds considerable detail which is not easily acquired by other means: "*We are interested in what happens to a patient.. it's really important to understand how the [medical] decisions were made. [GPs] are obviously thinking about [a lot] and they're taking these actions based on these things that they've noticed. We'll look what causes delay and wrong treatment, you've absolutely got to get that sort of detail. You can't have a questionnaire for that*" (E4).

Unlocking the richness remains a challenge, as the perspective of secondary analysts on data that originated in disjunct contexts, is 'hidden' in free-text fields or otherwise tacitly, silently contextual and by that evident to the involved co-workers, but beyond that 'lost in translation' (fig. 7.1).

In addition to their other epidemiological or HSR-related work, which mostly utilizes either (1) statistical data created by third parties or other intermediaries between health

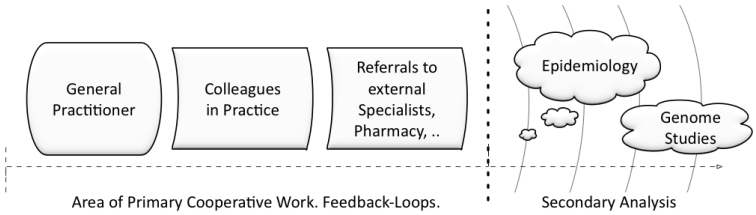


Figure 7.1: Usage of EHRs among primary actors (e.g. GP, colleagues, specialists, ..) is embedded in collaboration that, at least to some part, can rely on the „ground truth“ of patient care to negotiate and validate its process. Secondary Analysis is disjunct from this system and often has to make assumptions on how the collaborative artefacts that constitute EHR documentation are to be interpreted.

services providers and scientific research or (2) data collected primarily for clinical studies (E1, E2, E4, E5), the experts described the following sources for secondary analysis EHR data:

7.2.1 Reimbursement/Claims Related Data

For epidemiological studies which focused on medical data on the population of Austria, E1, E2 and E5 described the usage of the reimbursement/claims database that is run by the association of Austrian health insurance providers.
According to E1 and E2, a central quality issue with this data is its timeliness: Austrian GPs are reimbursed in quarterly periods, which requires for reimbursable services to be transmitted in the respective quarter, but does not hold a further incentive for GPs to document them more accurately, e.g. on the day of their occurrence. A data quality issue of no importance to the primary purpose of data collection, but problematic for efforts of secondary analysis with this data.
Furthermore, the interviewed experts who have experience with utilizing reimbursement / claims-related data, e.g. collected via the Austrian LKF[4] system, expressed different opinions on whether this data constitutes usable data for secondary analysis. While E5 described the usage of this data as challenging, but doable, E1 described it as a source that is "*practically worthless*".

7.2.2 Clinical Practice Research Datalink (CPRD)

E4 described a common source for primary care EHR data for purposes of secondary analysis in the United Kingdom: The Clinical Practice Research Datalink (CPRD). Formerly known as General Primary Care Database (GPRD), the CPRD covers 8,5% of the UK population with data from over 600 practices [58]; resulting in over 1500 publications since its introduction in 1987[5]. Data from the CPRD consists of patient-

[4]Leistungsorientierte Krankenanstalten-Finanzierung (LKF) - performance-oriented hospital financing.
[5]https://www.cprd.com/about/, accessed 1-Feb-2016.

related demographical information (patient's sex, age, year of birth and other details), each occurrence of illness or symptoms, all significant morbidity events (e.g. clinical consultations, diagnoses, test results, referrals and hospital admissions) [46]. "*The CPRD currently contains 4 billion records originally extracted from primary care practices, in 1/2 terabyte of relational data.*"[59]

7.2.3 Other

In conducting research in relation to influenza surveillance, E3 utilizes a **hybrid-form in between primary and secondary analysis**: Austrian GPs may participate in the Austrian Diagnostic Influenza Network (DINÖ)[6] and by that submit their EHR data by a side channel to E3 at the Viennese General Hospital.
In order to mitigate against data quality issues, the DINÖ demands from the participating GPs to conduct and document examination in cases of suspected Influenza-Like Illness[7] along a strict specification. E3 reported, that while the DINÖ-specification of the GPs examination of suspected ILIs is more thorough than general best practice / *lege artis* procedure, its documentation mostly does not differ from other Austrian GPs. Still, it would take participating GPs three to four years to reliably and consistently work in accordance with the DINÖ specification.

E1 described also **collaborating directly with Austrian GPs** who would approach Vienna Medical University, claiming to have data valuable for epidemiological research.

The strategies of approaching data described with these sources illustrate the challenge of obtaining valid and reliable data faced by secondary analysts. The following section 7.3 describes in more detail common notions of data quality and how secondary analysts validate them.

7.3 Dimensions of Data Quality and Strategies of Validation

7.3.1 Dimensions of Data Quality

Tate et al. remark, that "*although many books and articles have been written about data quality in general, there exists no commonly accepted methodology for defining and comparing data quality in medical databases used for research.*" [59] This section attempts to assemble a notion of viable EHR data for secondary analysis along the descriptions of the interviewed experts as well as several literature studies (fig. 7.2).

A EHR data quality relevant for secondary analysis and also for its primary purposes is **completeness**: Is an information present in the EHR? It is a consistently mentioned issue as well as the leading data quality problem in 4 of the 5 literature studies compared

[6] http://www.influenza.at, accessed 15-February-2016.

[7] http://www.who.int/influenza/resources/documents/WHO_Epidemiological_Influenza_Surveillance_Standards_2014.pdf, accessed: 15-February-2016.

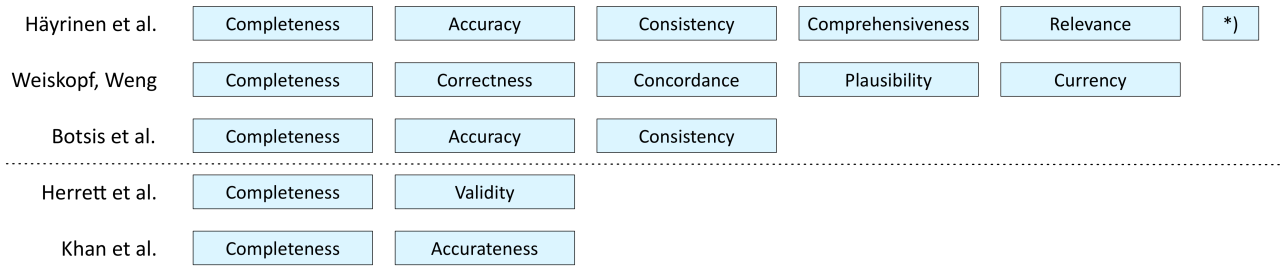


Figure 7.2: Different notions of data quality in Electronic Health Records, as found by Häyrinen et al. [47], Weiskopf and Weng [10], Botsis et al. [12], Herrett et al. [46] and Khan et al. [9]. The latter two focused on data quality for coded diagnoses, while the others on EHR data in more general respect. *).. Häyrinen et al. found further quality measures which other authors would, explicitly and implicitly, subsume mostly under *accuracy*.

in figure 7.2 [10][9][46][12].
Determining whether a given data set is **complete** can furthermore be more difficult than just checking for the presence of desired data fields, since the notion of completeness can be context-dependent, as described by Weiskopf et al.: Completeness of a given data set might differ depending on the object of interest (patient or health care process), the desired granularity (records as whole or components of records) or an intrinsic or extrinsic perspective (fitness of data in general or fitness in respect of a certain research interest) [13]. In extrinsic perspectives, Weiskopf et al. describe four categories of completeness (with more feasibly existing), depending on the research interest at hand: (1) **Documentation Completeness** denotes the presence of all observations made about a patient in the record, (2) **Breadth Completeness** denotes the presence of all observations desired by the research design, (3) **Density Completeness** denotes the presence of of data points in a desired frequency over time, (4) **Predictive Completeness** denotes the presence of sufficient information in the EHR so that desired statistical predictions are possible [13].

Another frequently described quality is **accuracy**. Weiskopf and Weng, using the term correctness (but explicitly capturing aspects of accuracy with it), describe it as the fact that a certain information is found to be *true* [10]. Botsis et al. broaden the definition towards information being *specific*, *exact*, and/or *precise* [12]. Several different types of record qualities can be considered to be adding to the record's accuracy; e.g. by Häyrinen et al., who identified traits of quality as distinct, which other researchers would subsume under the concept of accuracy: *timeliness*, *relevance* and *format-abidance* [47].

Further identified qualities include **consistency** (Weiskopf and Weng defined it by the question of *Is information present in the EHR coherent to/in agreement with other data sources*?), **plausibility** (*Does information make sense in relation to other knowledge and expectations about the element?*) and **currency** (*Is the information relevant to represent the patient state at the time of its creation?*) [10].

These quality factors were also described consistently by the interviewed experts of

epidemiology and HSR, with completeness and accuracy being most important, as well as a general notion of data having to *make sense* (E1). "*I expect clinicians to remove pointless data, which they do only partially. If I see, that data jumps around extremely, it looks suspicious to me, but I cannot estimate real quality, as long as its heritage is not completely, logically comprehensible to me.*" (E2)

The Interest in Structure

Apart from the qualities described in the previous part of this section, secondary analysis of EHR data exhibits affinity for structured data, as such data lends itself better to statistical analysis. In an evaluation of structuring in clinical records conducted for the NHS, Morrison et al. acknowledge a "*rise in expectations and emphasis on the use of structured and/or coded data for research*" and the "*strategic interest in increasing the proportion of the EHR that is captured in structured and coded format*" [22].
While Morrison et al. state, that imposing structure on EHRs "*generates unambiguous data which improves accuracy in secondary uses*" and that "*there was clear and consistent evidence of substantial opportunities arising from the secondary uses of coded data for public health, healthcare management, audit and research*", they also remark, that the question of how the facilitation, maybe even enforcing, of structure and/or coding in EHR usage practice might affect the viability of the resulting data, including but not exclusively for secondary use, still lacks empirical investigation [22].

In addition to its perceived value for secondary analysis, promoting and facilitating more structure and coding in EHRs is also driven by (and vice-versa, drives) objectives of improving interoperability between different EHR systems, networks and databases. Tolar and Balka name the development of a uniform, standardized terminology as a "major barrier" for the development of interoperable infrastructure [57], since variations in semantic representations among different EHR systems are common [12].

To the interviewed experts, the presence of structure and codes is an important quality factor for the secondary analysis of data emerging from the clinical routine. "*Qualitative features are difficult to analyze, so I usually give that back to the clinician and say: 'Go make something codeable out of that'; typically ICD9 or ICD10-codes.*" (E2).
Structure has its ambivalence, as NLP-Researcher E6 remarks: "*The more structured an information is, the better I can search. On the other hand, users are limited by it. They want to write down something, and then think to themselves: 'This doesn't quite fit there. It's a trade-off*" (E6).

7.3.2 Strategies of Validation

In relation to the dimensions of data quality of interest to secondary analysts described in section 7.3.1, researchers and facilitators employ various methods of assessing the data in order to ensure its quality. This section describes common strategies of validation as described in literature and by the interviewed experts. In line with Boslaugh's three questions for assessing data [8] for the purposes of secondary analysis mentioned in

the beginning of this chapter, taking on general practice EHR data necessitates the familiarization with its peculiarities, strengths and weaknesses, and efforts of validation will vary depending on this.

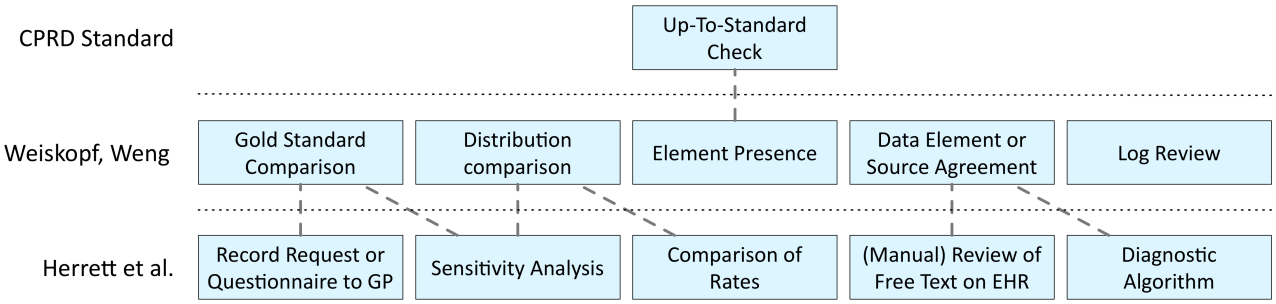


Figure 7.3: Common validation strategies with EHR data. CPRD-Standards intrinsic to the database (top), Strategies of validating EHR content in general [10] (middle), Strategies to validate diagnoses [46] (bottom). Dotted lines represent conceptual relationship.

7.3.2.1 Intrinsic Fitness By Example of the CPRD

For secondary analysis of EHR data, various intermediaries exist who collect, process and provide this data to researchers, like the organizations CPRD[8], THIN[9] or DIN [60], who provide such databases on EHR data originating from the United Kingdom's health system. In Austria, the country five of the interviewed experts are based at, the provision of EHR data for secondary analysis purposes is offered by Statistik Austria[10].
Providers of such databases conduct intrinsic fitness evaluations and uphold criteria of inclusion in order to establish a baseline data quality for subsequent further uses. In Statistik Austria, a medical statistics division defines and enforces such criteria (E2). This section describes such criteria as installed in the UK's CPRD as an example; it is believed to be more relevant to the thesis overall, since it covers the UK health care system and thus relates more to the three EHR systems investigated in the previous chapters 4 to 6.

The data fitness assessment for the CPRD includes a practice-related 'up-to-standard' (UTC) check described in the quote below, as well as patient-related checks:
Individual patients are flagged as unacceptable for use in secondary analyses in case of 'non-contiguous follow-up' or significantly incomplete data recording [61].

> "The [CPRD] carries out a series of on-going checks to ensure that the data are 'up to standard'; this comprises assessment of both patient data (age, gender, registration details and event dates) and the completeness, continuity and plausibility of electronic data recording in key areas at the practice level (for

[8] www.cprd.com, accessed 15-February-2016.
[9] www.ucl.ac.uk/pcph/research-groups-themes/thin-pub/database, accessed 15-February-2016.
[10] www.statistik.at, accessed 15-February-2016.

> example, ensuring a minimum specified percentage of deaths have cause of death recorded, a minimum referral rate per 100 patients, and a minimum number of prescriptions per patient per month)" [46, p. 5].

This approach to the data's validation succeeds in providing evenly complete data for further research purposes, but has its limitations as well. Williams et al. stated in 2012, that "*no specific validation work has been conducted on this method as much of it is based on logical inconsistency of the registration data*" [61]. Herrett et al. speculate, that this up-to-standard requirement of acceptance, while on one hand providing better-quality data, also has an effect on the data's generalizability; practice contributing to the CPRD are more "*likely to record disease events better than general practices that do not contribute*" [62]. A question raised also by Thomas et al. [63]

As Khan et al. remark: "*While the checks conducted by the GPRD team may provide an overall evaluation of which practices are providing good-quality data meeting certain standards, they do not specifically assess the validity and completeness of individual patient records. [...] As the GPRD is increasingly used for academic research, it is important to consider the quality of the data available for study*".[9]

Beyond validating the intrinsic fitness of data aggregation databases as described above, several approaches to the validation of their content were described by the interviewed experts as well as are described in related literature. Figure 7.3 illustrates two further levels of validation: Weiskopf and Weng [10] reviewed literature to identify strategies of validating EHR content in general. Herrett et al. [46] reviewed literature to find such strategies of validating specifically diagnostic information. The difference in perspective appears negligible, although certain approaches become only feasible under specific validation questions, e.g. reviewing the free text of an EHR documentation manually.

7.3.2.2 External Comparative Validation Approaches

Several described approaches compare the EHR information in question with external data sources, which can be taken on as a point of comparison or a source considered benchmark-valid *a priori*, commonly referred to as **Gold Standard**. Commonly sources for gold standard information are information supplied by patients, data reviews with patients [46], *paper records* kept by the GP or *questionnaires* [10]. The selection of gold standard sources has to be made carefully, as Herrett et al. remark: "*Practices who do participate in validation studies may only send information for certain cases, e.g. refusing to copy very large case files; this may result in selection bias*" [46].
Other forms of external comparison include **distribution comparison**, where statistical features of the data to be validated are compared with expected distributions, e.g. with that "*of a restricted set more likely to represent true cases*" (**Sensitivity Analysis**) [46] or with distributions of other, non-CPRD, UK-based data sources (**Comparison of Rates**) [10].

7.3.2.3 Internal Comparative Validation Approaches

Furthermore, comparison with data sources internal to the EHR documents is conducted to establish their validity. This can entail checking several different data entries of an EHR document on whether they are plausible to one another (**Data Element Agreement**), e.g., in the case of diagnosis-validation, by *manually reviewing the free text* associated with the EHR document for explicit or implicit information in regard to the diagnosis in question [10].
Drawing conclusions from constellations internal to EHR documentations within a database is also done algorithmically, e.g. **Diagnostic Algorithms**. Algorithmic approaches can help capture complexity caused also by the ambiguity of the coding regimes and its different possibilities to code a medical problem. Morley et al. developed a diagnostic algorithm to identify patients of atrial fibrillation (AF), a chronic condition, and found 201 Read codes to be used in documenting AF diagnoses [64]. Diagnostic algorithms sometimes lack transparency in scientific publication, as Weiskopf and Weng remark: "The medical codes used to define cases were seldom reported in papers, although some researchers indicated that they would make code lists available" [46], a circumstance complicating the reproducibility of the results. They furthermore describe possible disadvantage of hidden biases in selection: "*[The] use of such algorithms may exclude less severe cases that do not require treatment, and the inclusion of test results in these algorithms is problematic since not all test results are recorded in the [CPRD].*"[46]
Algorithmic approaches are also used to conduct probabilistic estimations. Nicholson et al. used not diagnosis-related codes and statistically analyzed their co-occurrence with codes denoting the diagnosis of rheumatoid arthritis, concluding that "*a broader range of codes which may allow a probabilistic case definition to use in cases where no diagnostic code is yet recorded*" [15].
E6 remarked, that algorithmic approaches might assist with making causal and correlational connections where none would otherwise have been suspected and thus also not investigated. Such systems are subject to intense research to the end of providing doctors and other medical personnel effective ways of analyzing a growing amount of patient-related data. E6 mentioned an example of algorithms being designed to detect cancer-diagnoses of the parietal lobe even when not explicitly described by these words.

Manual approaches of validation, e.g. by manually reviewing free text portions of an EHR document, hold reliability in that there are still many complications with automated reasoning and understanding of EHRs, as E1 remarked from his experience: "*I struggle with letting a machine do [validation]. Medicine is full of synonyms, also many results and accompanying documents are contained within the record which are not included in automated assessments*" (E1).

7.3.2.4 Others

Other categories of validation approaches mentioned by Weiskopf and Weng are checking for *element presence* in order to ensure the respectively analysis-relevant form of completeness, and *log review* - drawing conclusions from "actual data entry practices (e.g.

dates, times, edits) [10].

7.4 Phenomena of Variability

In order to sustainably improve data quality in respect of the qualities in EHR data desired by secondary analysts as described in section 7.3, a detailed picture of the data's heritage, creation and processing procedures is necessary. Identifying and understanding causes of variability in various aspects of data quality thus is of key importance to data validation for purposes of secondary analysis.
This section discusses types of variability in EHR data as unearthed in the PREP field study's analysis process and discussed in the expert interviews, related to literature.

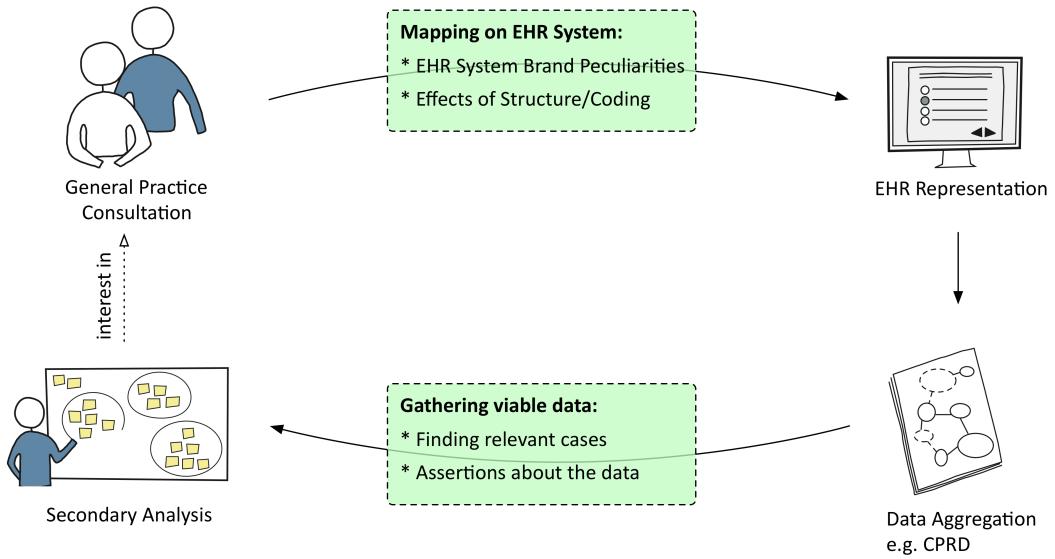


Figure 7.4: Flow of data from its origin in general practice consultations over data aggregation efforts to secondary analysts.

7.4.1 Experts' Expectations of PREP Results

After an explanation of the study's setup, the experts interviewed in the course of this thesis described their expectations of the PREP field study's results differently. All experts expected that there would be at least some variability in consultation outcome, e.g.: "*I would expect results to be very diverse, because some would probably take the safe route with antibiotics, so to say, and others would try to avoid that at first. You know, some physicians might give something hard, so that they are not bothered anymore, others take a less intense route, try some alternative methods*" (E5). "*I would expect, in the joint pain case, the diagnosis to be more variable than with the throat. They can look into the throat, but with inflamed joints he can hardly do that, then he'll be confused*" (E2). E1 described this as an aspect natural to primary care: "*Without doubt, results*

will vary. This is a basic problem in epidemiology: We don't have diagnostics in primary care, only comparison with 'images' of diseases" (E1).

At least some variability would be expected also in EHR documentation. Generally, most of them mistrust coded data as a source for reliable information from primary care, but find it usable in epidemiology given sufficient validation: "*Generally, I expect little of data coming from routine work. We already know from experience, that these diagnoses that are officially reported, they are.. probably randomly generated or something like that, because they don't.. because it's about money. So we don't trust routine data*" (E2). An opinion also shared by E1, E3, and E6. "*I don't mind analyzing routine data. As long as the people who draw the conclusions are aware of the fact, which criteria these data fulfill and don't fulfill*" (E1).
E4 holds GPs in significantly higher favor: "*If you talk to a specialist, they tend to be quite critical of primary care records, they tend to say they are not very detailed and this kind of things.. and that you should get much more precise coding. If the world were perfect, you would have everything perfectly coded and there wouldn't be any extra information. I think that there is generally underappreciation of the complex range of things that people are trying to do when creating records, and so if I write some stuff that in code, if I give a non-specific code like painful joints, and then I give some other information in the free text, any doctor will think, ok this is almost certainly RA waiting for formal diagnosis. Then that is there. It is there in the information. So you can't say it is not there because it has not been properly coded* " (E4).

In terms of suspected, underlying reasons for variability in results, the experts pointed towards several aspects described in more detail in this section: (1) Incomplete records (s. 7.4.2), (2) variability in codes in general practice (s. 7.4.3), (3) variability among databases (s. 7.4.4) and (4) variability caused by differences in system design (s. 7.4.5).

7.4.2 Incomplete Records

As described in section 7.3, incompleteness in EHR data can have many faces, depending on the underlying requirements of completeness, and represents a very common data quality issue. Incompleteness can have various underlying reasons and, as Pivivarov and Elhadad remark, evidence is mounting, that most of the missing data in the clinical domain is *Missing Not at Random (MNAR)*:

> For these data, the missingness is informative, meaning that there is an underlying reason that the data are missing but that this reason is simply unobserved. [65, p. 943]

Finding and understanding reasons for MNAR-incomplete EHR data can improve mitigation of incompleteness by statistical means: Unlike MCAR (*Missing Completely At Random*)-relations, denoting complete randomness in the presence/absence of a data element, or MAR (*Missing At Random*)-relations, denoting that the absence of a data element can be predicted by a constellation in other data fields of the data set (but not

its content), MNAR-relations also allow for assumptions on the content of missing data element. Lin and Haug explain [66] this by the following example:

> A physician may assess the lung's condition by subjective complaints and auscultation. These variables are not present in the sample data. A chest X-ray may not be considered necessary if the physician feels the patient's lungs are normal and, in most cases, this inference will be correct. The absence of chest X-ray does not depend on the observed data in the data set, but on the missing chest X-ray value guessed by the physician using some mechanism not reflected in the data. [66, p. 3]

Instances of comparative incompleteness were found in the PREP field study in comparison among surgeries as well as used EHR systems: Among others: (1) Different interaction design of data entry interfaces causes variability in structure and coding and thus presence of information in many further statistical analyses, and (2), notably: GPs from Surgery HS described their practice of not issuing a Rheumatoid Factor (RF) blood test as per the wishes of the rheumatoid arthritis specialist resident in their area (see chapter 4). This constitutes valuable insight into the presence and absence of RF blood test data in aggregated EHR data. Sammon et al. found RF test results to be MNAR in the CPRD database, hypothesizing that, because negative RF results are of less clinical relevance, they do not get reported [67]. The rationale of HS surgery's GPs may point to another reason for missing RF test results and question these implications drawn from the MNAR-relation.

Incompleteness was consistently reported also by the interviewed experts as the prevalent quality problem with EHR data for purposes of secondary analysis. E4 remarked, that time constraints in GP surgeries require GPs balancing the time they devote to documentation, causing 'low-stakes' medications, while being very interesting information for secondary analysis, to sometimes not be documented; this was also observed in the PREP field study's analysis with over-the-counter medication (see chapters 4 and 5). E4 and E2 furthermore remarked, that missing structure in EHR system infrastructure causes incompleteness as well. While blood-test results are automatically documented in a structured fashion, no such automation exists for x-ray results, causing them to be mentioned in free text, if at all (E4).

7.4.3 Variability in Codes in General Practice

A considerable amount of variability in EHR data, coming from general practice as well as generally clinical EHR data, originates from the usage of codes to document information and can have different underlying reasons. Imposing a specific vocabulary on the documentation of medical work constitutes a comparably strict form of structure and the ability of such coding vocabularies to capture information in a meaningful, expressive way is disputed, as the expert interviews revealed.

This is also substantiated by the PREP field study analysis results; as GPs predominantly considered the usage of codes in their work as a useful measure influenced by motivations to improve searchability and document their work for reimbursement or QOF purposes (see chapter 6). Furthermore, GPs considered codes to have the character of an official, labeling action, letting them hesitate with applying codes that might 'expose' their patients, as described in chapter 6, in line with findings of Hardstone et al. [4].
The Read code system, as described by GPs, furthermore exhibits ambiguity and redundancy in its code vocabulary, as is also reflected in literature: In a study on the usage of antibiotics in sore throat cases, Dunn et al. found that "*the diagnostic label attached to a patient with any RTI will vary with different healthcare professionals*" [53]. Morley et al., in their study on disease phenotypes of Atrial Fibrillation (AF), remarked, that they "*found, as have others, that although codes for AF subtypes exist within the Read system, they are infrequently used and most patients simply have an all-encompassing diagnosis of AF recorded.*" [64].

This section details on two aspects with code practice that emanate variability in coded EHR data: Incentivization by reimbursement processes (7.4.3.1) and the viability of coding in relation to the work of GPs (7.4.3.2).

7.4.3.1 Incentivization By Reimbursement Processes

The effect of pay-for-performance incentives like the Quality and Outcomes Framework (QOF) on general practice has been shown e.g. by Vamos et al., who described the effectiveness of QOF in pushing the realization of UK national targets for diabetes management [68] and Chew-Graham et al., who conducted a qualitative study on the practice and perception of QOF and found it to be a powerful, shaping force on general practice [69], as did Williams et al., finding that recording practices have "changed enormously" with QOF rules. [61]. Mathur et al. compared the completeness of ethnicity recording within the CPRD around QOF incentivization between 2006 and 2012 and concluded that this "dramatically improved the completeness of ethnicity data for newly registered patients" [70].

This powerful incentive to using coded information especially when favorable to the GP's (or rather their employing organization's) financial reimbursement is seen as a strong factor of bias in EHR data; with "uncertainty surrounding diagnostic accuracy in administrative claims" [71], the validity for secondary analysis purposes lies in question. In the PREP field study analysis, QOF - e.g. "hunting for [QOF] targets" - was mentioned frequently by GPs when describing influences on their coding practice.

To the interviewed experts, influences by reimbursement processes on the resulting EHR documentation are well known also from their experience with the Austrian health care system: "*Insurance is a disturbing factor. I don't insinuate general practitioners acting with ill intent, but I just know, that accounting influences the documentation, there is no doubt about that. I can still take it for analysis, but I have to point to the fact, that this factor of distortion is there. [..] In another area, the hospital sector, we know for a*

fact, that the LKF[11]*-system is bent in a way to score the most points for the hospitals. Computer programs exist, that suggest for every single patient which examinations to do, because they bring in points. As soon as economics is in the equation, things will be documented that are not exactly true*" (E1). A concern shared in almost similar words by E2 and E5.

7.4.3.2 Coding vs. Nuance - Viability of Codes in GP Work

Given the influencing factors and mere chance that surround the usage of codes in GP EHR documentation, the question arises whether coding vocabularies are at all able to capture the information of GP consultations in a meaningful way. As Morrison et al. found in their evaluation of UK EHR systems:

> Clinicians tended to find it hard to choose an appropriate coding category, often picking the closest match as opposed to an accurate description of the situation. This had important accuracy implications for secondary users of the data. Most healthcare professionals felt detached from the development of codes, which possibly resulted in reduced clinical utility. [22, p. 50]

The necessary task of mitigating against factors of chance and uncertainty in GP documentation is faced by the interviewed experts in different ways. While the strategy of epidemiologist E3, working on influenza surveillance - ensuring code validity by enacting stricter rules, described in section 7.2.3 - might prove effective for the specific task at hand, it stands to reason, that scaling it up (1) thematically beyond influenza to all areas of GP work and (2) beyond 45 participating GPs who can be supported individually to all GPs of a country or area might not work, given the fact that even under these focused circumstances, it would take GPs three to four years to produce reliable data, as E3 reported. A more viable approach seems to be required in order to better handle the mapping of nuance, as it naturally occurs in the GP profession, in EHR documentation.

Turning to 'Better' Codes

Diagnostic codes, as E1 describes, exhibit significant variability because of the fact, that they do not adequately reflect the practice of medicine in the field of general practice and thus do not easily and unambiguously lend themselves to document this work: "*In our study, we asked GPs to document codes related to medical complaints only. Then, we had two independent GPs interpret these for diagnoses, looking for how well they would match. They matched at roughly 2/3 of cases.*" (E1), implying also that the vocabulary of ICD-10 diagnoses does not yield an always accurate mapping of medical complaints in primary care.
E1 described such circumstances as natural to the profession of general practice: *"This is a fundamental problem in epidemiology, that in primary care, we do not have diagnostics, but usually a more or less well trained, more or less static to the state-of-the-art of*

[11] Leistungsorientierte Krankenanstalten-Finanzierung (LKF) - performance-oriented hospital financing.

their training at the time, physician who tries to get a picture of a medical problem. Because, in many cases, these physicians do not have the diagnostic means, they have to resort to comparison with 'images' of disease. Furthermore, the work of primary care is incremental. Braun[12]*, the father of scientific primary care, said: 'The common is common'. At first, GPs will treat common causes, then, if this won't help, further diagnostic measures, for example in a hospital, will follow. But in the common case, this does not happen, and the diagnostic remains at images of disease. 'Looks like', 'has complaints as if', this can not be expressed by most codes, e.g. ICD-10, so I need different ones."* (E1). An assertion confirmed e.g. by Fink and Haidinger, who, in a statistical analysis of 10 years of health record documentation in an Austrian GP surgery of one GP, found, that "49% of the episodes of care had to be attributed to 'symptoms and groups of symptoms' (without any confirmed diagnosis), 40% to 'picture of a classic disease, but not confirmed' and only 11% were 'definite diagnoses' (confirmed)" [73].

Subsequently, E1 turned to ICPC-2 (International Classification of Primary Care, Second Edition), a coding vocabulary accredited by the World Health Organization (WHO) for use in primary care. According to E1, using ICPC-2 would help mitigating against reimbursement-related bias as well as better tailor to the needs and practice of general practice medicine.

Turning to Free Text

Another approach taken in mitigating against challenges in meaning-accuracy with codes is turning to free text portions of EHR documents for means of validation and general information extraction. While on one hand, this approach works hand in hand with improving coding vocabulary, on the other hand, free text holds information that, according to E4, can not be coded easily and might not be codeable at all: "*Thinking that you just have to find the right code, and then you don't need to say anything to anybody - I mean this is just, this fundamental vision on what language is, there just is no evidence of that. If you look at the kind of variety people need to write, these are human interactions. The Question is, what do people need to know, what do they need to agree between them while they're there, what do they need to pass on to other people. Differences might be that they might see that person later. There is value in the facts of diagnoses, but that doesn't exhaust what you can get out of the records, because they clearly tried to do a lot more than that. [..] The fact, that computational linguistics cannot today pull out all the detail of things that we would like to know about doesn't mean there's anything wrong with how we're recording.*" (E4).

A core advantage of free text in its approachability is that it does impose none or far less structure on the person attempting to convey information by it. As E4 remarks: "*There are lots of things you might want to talk about in the text, that you're not making statements about. [..] What doctors are doing is showing the next person they've considered certain possibilities, which, for instance, are not true and not relevant, and hence they're*

[12]Robert N. Braun, 1914-2007 [72]. http://www.oegam.at/wissenschaft-publikationen/aus-der-allgemeinmedizin/robert-n-braun/, accessed 15-February-2016.

not doing this x-ray or that text, and so there's a huge amount of .. demonstrating your train of thought that can never be captured by coding, because codes are statements as 'X' or 'not X'. They are logical operators. They might be boolean operators, AND, OR, .. but they cannot capture any of that" (E4).

The notion that the free text constitutes a valuable source of information is confirmed by studies [15], e.g. as information that can improve diagnostic algorithms [74] or make more detailed information visible prior to the presence of definitive, coded diagnoses, as e.g. shown by Tate et al., who analyzed free text and found that, in cases of ovarian cancer, 22% of cases had free text confirming diagnoses prior to the date of coding. [14]:

> There is no Read code for a possible, probably or highly likely diagnosis, and this may explain why 10% of the patients classified as having only a suspected or ambiguous diagnosis nevertheless had an ovarian cancer code. [14, p. 12]

Two possible approaches to mining free text are (1) manual annotation of free text by GPs or professional annotators, as described e.g. by Koeling et al. [75], and (2) employment of Natural Language Processing (NLP) to automatically retrieve information.

Since free text is not in the focus of this thesis, further description of free text practices and approaches are beyond the scale of this work and no further detail will be mentioned here, albeit available in e.g. the works of Rijnbeek [76] and Williams et al. [61].

7.4.4 CPRD in Comparison to Other EHR Databases

Khan et al., in their literature review of 46 publications reporting on secondary analysis on the basis of CPRD data and their validation efforts with Gold Standard strategies, concluded that most of the diagnoses coded in the CPRD exhibited good validity [9]. They furthermore remark:

> Generally, there is good agreement in disease prevalence rates between the GPRD and other national databases and statistics. [..] There is no 'gold standard' measure against which data from one database can be compared, or to suggest which database contains the most accurate measure.[9, p. 133]

The unavailability of a viable Gold Standard to compare against makes the identification of underlying reasons for differences among databases an important aspect in database validation. In comparison with other databases of EHR data, e.g. the Doctors' Independent Network (DIN) database or the The Health Improvement Network (THIN) database, exhibited differences, some of which not exhaustively explained.

CPRD/DIN: Springate et al. repeated a study conducted on the basis of data from the DIN with data from the CPRD and found, that

> [d]espite purporting to cover the same general UK population, there were some notable demographic and clinical differences between the Clinical Practice Research Datalink and Doctors' Independent Network cancer cohorts. Sensitivity analysis indicated that these had only a minimal effect on treatment effect estimates, but we were unable to account for a difference in mortality rates between the cohorts. [77, p. 1]

Notably, they remark, that both studies point to "*very different conclusions regarding the safety of beta-blockers for solid patients with cancer*" and strongly recommend to validate studies using independent, additional data sources [77, p. 9].

CPRD/THIN: In a study on the estimation of burden of disease with venous leg ulceration conducted with data from the CPRD and THIN databases from 1998 to 2006, Petherick et al. found that, while differences were significant between 1998 and 2000, results were comparable from 2000 to 2006 [78], speculating that the decreasing differences are caused by growing technology literacy and abidance to database fitness requirements:

> Irrespective of the primary care database that they contribute data to, new practices are more likely provide incomplete data as they learn to use new computer systems and achieve new quality standards of clinical data reporting. [..] During the study period investigated, more practices have joined the THIN database, including half of those that also contribute to the GPRD, while in contrast more practices have stopped contributing to the GPRD. By the year 2000, more of those practices that joined the THIN database had contributed data electronically for several years and had met the acceptable mortality reporting standard required by the database. [78, p. 378]

7.4.5 Variability Caused by Differences in System Design

As became apparent in the PREP field study's analysis (chapters 4-6), the three different EHR systems at hand (Vision, SystmOne, EMIS LV) exhibit various differences in interaction and interface design, which also affected coding and documentation activities. Among the interviewed experts, these types of variability phenomena were anticipated differently:

While E1, E2 and E4 are aware of differences in design of code and free text entry as well as other aspects across EHR systems, only E4 would expect this to affect the data available to the secondary researcher: "*Clearly there is quite a big difference with the software systems, which is not legislated for in any way by the standards of what is an adequate software system*" (E4). E2 expects medicine statistics to mitigate against variability emanating from differences in EHR system design. E5 would not expect the interface design to matter and better be left open for personalization and differentiation based on "*tastes and needs*" (E5), while data structure harmonization across systems is

important, but only an internal matter: "*A defined data structure is necessary, without it, interoperability of systems will not be possible. But how this loos like, what is done with it, is not that important, I think. This should be left to the EHR system creators, after all there are differences and physicians are used to a certain interface and structure. They usually don't want to switch. I wouldn't try to harmonize this locally, because there are different approaches, different needs*" (E5).

Williams et al. would suggest that differences between EMIS, SystmOne and Vision exist, as "the system design encourages slightly different recording styles", but the "*implications for research [..] are likely to be minimal*" and can be controlled for [61, p. 90].

This aspect is further explored in chapter 8 (*Discussion*).

7.5 Summary

Chapter 7 summarized the results of a literature review and expert interviews conducted in the course of this thesis, focusing on aspects of secondary analysis of primary care EHR data. It described the necessity of careful consideration in handling data for purposes of secondary analysis in order for it to yield meaningful results. Section 7.2 described common sources for primary care EHR data in secondary analysis. Section 7.3 gave an overview of common notions of data quality as well as validation strategies. Section 7.4 described phenomena of variability as exemplary instances of underlying reasons of variability that may affect data validity for secondary analysis. Validity of primary care EHR data for purposes of secondary analysis is a multifaceted, context dependent notion, that can be affected by various phenomena of variability on different levels.

The PREP field study's analysis could describe several causes of variability. In part, these results confirmed prior results in literature and/or experiences described by the interviewed experts:

- In primary care EHR documentation, codes are considered not trustworthy and in need of thorough validation; results of subsequent secondary analyses require a considerate statement of related limitations. Free text is considered to hold valuable information, but is hard to access systematically and automatedly.
- Incompleteness is a commonly mentioned problem with different causes and circumstances depending on the underlying definition of completeness and the type of incompleteness (missing at random, missing completely at random, not missing at random).

Phenomena of variability that originate from peculiarities of the employed EHR systems in terms of interface and interaction design are known to exist, but believed to play only a marginal role in causing variability in the resulting EHR data. This aspect is taken on for more a thorough discussion in chapter 8.

CHAPTER 8

Discussion and Conclusion

This chapter constitutes a discussion of the results of this master's thesis work and a conclusion on the basis of this discussion. Section 8.1 (*Overview of Results*) summarizes the results of the individual approaches taken in this thesis and relates them to the research questions. Section 8.2 (*Contribution of this Work*) describes the contribution the results of this thesis yield to the field of EHR system design. Section 8.3 (*Implications for Design*) formulate implications for the design of EHR systems as part of this thesis' contribution. Section 8.4 (*Limitations*) describes limitations that need to be taken into account in relation to this thesis' contribution. Section 8.5 (*Conclusion and Future Work*) ends this thesis with a short conclusion statement and pointers to possible future work.

8.1 Overview of Results

In the course of this master's thesis, data gathered in a field study in relation to the PREP project was analyzed. This analysis was driven by two research questions that were put in place in order to identify possible forms of variability in consultations and EHR documentation across the three different brands of EHR system employed (Vision, SystmOne and EMIS LV) as well as across the surgeries:

- RQ 1: *How are the two simulated scenarios dealt with in the consultations and how is the information originating from the consultations documented in the EHR systems at hand?*
- RQ 2: *How does this relate to the understanding GPs have of the building blocks of EHR documentation - coded entries and free text?*

To this end, the three different perspectives on the PREP field study - (1) video recordings of the simulated consultations with 15 GPs in 5 surgeries, (2) screen recordings of EHR

usage and (3) walkthrough interviews with the participating GPs - were taken into account to form a rich picture of the EHR-system-related practice of primary care physicians in medical consultations, as described in chapters 4, 5 and 6: Chapter 4 describes results found in the course of the video analysis of (1). Chapter 5 focuses mostly on (2), taking into account relevant views and accounts of experience given by the GPs in (3). Chapter 6 draws a picture of EHR system usage practice based on (3).

Furthermore, expert interviews were conducted with 4 epidemiologists and 2 experts on EHR design and NLP, respectively, backed by a review of related literature. This was done in order to relate the PREP field study analysis results to aspects of secondary analysis:

- RQ 3: How is EHR data coming from general practice seen by epidemiologists conducting secondary analysis with this data?

This resulted in a detailed view on practical aspects and challenges described in chapter 7.

8.1.1 Compilation of PREP Field Study Analysis Key Take-Aways

For purposes of clarity and comprehensibility, the key take-aways of each analysis-related chapter are compiled in this section as they can be found in sections 4.5, 5.3 and 6.4, respectively.

8.1.1.1 Research Question 1a: *How are the simulated scenarios dealt with in the consultations?*

In order to answer research question 1a, chapter 4 can be used. The simulated scenarios allowed for a rich insight in how GPs structure consultations, conduct their medical reasoning and document it in the EHR system:

Key Take-Aways of Chapter 4 / Video Analysis

1) The simulated patients were, to the most part, enacted consistently and invoked homogeneous medical responses by the GPs. They were consistently described as credible and the resulting consultations as realistic by the GPs.

2) While some variability was observable, the **consultations presented themselves as overall similar in their course and outcomes** across GPs and surgeries.

The analysis hints on experience and prevailing opinions on the side of the GPs, influencing the reasoning process in the construction of an understanding of the problem and the formulation of a plan, causing variability in how the cases of *Jane* and *Scott* were addressed by them. These differences would unfold on a per-case level: *Scott*'s case divided the GPs slightly along the question on to what degree the construction of a problem-understanding would narrowly focus on the observable signs (a sore throat) or

also take his environmental stress factors into account. This would affect what further measures would be prescribed by the GP. However generally, GPs would follow *Scott*'s enactment of putting main focus on his environmental stress rather than the sore throat problem, interpreting his case as more of a social/psychological problem than the effects of an infectious disease. *Jane*'s case divided the GPs along the question on to what degree her case would warrant for "urgency". This would cause some variability in the plan formulation phase, but would mostly play out implicitly: In GPs crafting the plan towards a swift referral to a rheumatologist.

3) Notably, **regional differences can influence the approach GPs** take to cases of suspected rheumatoid arthritis, as would be hinted on by the blood-tests issued for *Jane* in surgery HS and the influence on them by the rheumatologist local to the area of the surgery.

4) Some **further aspects of variability** were observable around the structuring of the consultation, in that younger GPs would rather adhere to a more step-by-step course through the basic focal theme stages described in section 4.3, while older, more experienced GPs would tend to an iterative approach, e.g. switching back and forth between the establishment of the present problem and the exploration of context, refining the questions with each iteration. Furthermore, the approach of GPs would exhibit variabilities in to what degree they would involve the patient in the decision making process.

8.1.1.2 Research Question 1b: *How is the information originating from the consultations documented in the EHR systems at hand?*

In order to answer research question 1b, chapter 5 can be used. The comparative content analysis revealed a variety of differences across EHR systems:

Key Take-Aways of Chapter 5 / Comparative Content Analysis

1) In their approach to assisting GP work, **EMIS LV, SystmOne and Vision structure and display information and integrate Read codes differently**. This encompasses different notions of a) What categories are available to the GP, b) How GPs can enter Read codes and c) How free text an Read codes relate to one another.

2) EMIS LV, SystmOne and Vision, according to GPs, come with **different advantages and disadvantages in usage practice** and they feature **peculiarities** that entail a necessity for different skills in GPs. For example: Vision's code-centric design requires GPs to know a specific Read code vocabulary. SystmOne's code prediction feature is challenging for GPs who are bad at touch typing.

3) Content analysis of the EHR documentation from the consultations with *Jane* and *Scott* revealed that the differences across EHR systems resulted in system-related **differences in the resulting usage of Read codes and structuring of consultation documentation in general**.

4) A comparison across surgeries hinted on **different norms** on documenting consultations with EHR systems **across surgeries**. For example: GPs from surgery HS are less verbose in their reports than others. GPs from surgery EH use less categories in their reports than others. This is described in more detail in the following chapter 6.

8.1.1.3 Research Question 2: *How does this relate to the understanding GPs have of codes and free text?*

In order to answer research question 2, chapter 6 can be used. The medical consultations with the simulated patients provided an entry point for the GPs to reflect on their practice; the walkthrough interviews yielded a rich picture of EHR-system-related practice of the participating GPs

Key Take-Aways of Chapter 6 / Walkthrough Interviews

1) The usage of Read codes is incentivized by several aspects of **usefulness** that are not directly medically relevant to the GPs' work, causing **differences in the likelihood certain information might be coded** with Read codes: (a) Information especially valuable for (a) QOF purposes or (b) structuring a medical record might be more likely to be coded.

2) Read codes might have the character of a label for a patient, causing GPs not to use codes for information that is considered very relevant, but also uncertain, as they find it compromises doctor patient confidentiality and might have unwanted ramifications.

3) The Read code database is extensive, containing lots of subtle differentiations that are considered unnecessary redundancies from the perspective of the work a GP does. Finding the 'right' Read code can be **influenced by chance**.

4) Free text is used to document **important** information that assists with memory, with reasoning, that acts as security for medical-legal purposes and supports cooperative work.

5) In GP practice, **EHR systems are** used in a field of tension between the necessity to document information and do work on one hand and the time this takes away either from doctor-patient rapport during consultations and from GP work in general, **exerting time pressure**.

6) When working with EHR systems, GPs bring their experiences and viewpoints to the table. They **continuously develop and change their EHR work practice** alongside experience and the exchange among surgery co-workers. In certain aspects, they **align towards common surgery-understandings of 'good' EHR-keeping**.

7) GPs reported, that, under certain circumstances, **EHRs can be retrospectively changed** to adhere to requirements of automation or because they look messy. GP14 furthermore recounted a situation of retrospective change of EHR data en masse for QOF reasons.

8) When creating an EHR entry, **GPs are mindful of the patient's view** on EHR documentation, **weighing the necessity to document certain information against possible ramifications for their relationship with the patients.**

8.1.2 RQ3: How is Primary Care EHR Data Seen By Experts and What Can PREP Results Contribute?

As described in section 7.5 (summary of chapter 7 on *Secondary Analysis of Primary Care EHR Data*), such secondary analysis of primary care EHR data requires a detailed understanding of the strengths and weaknesses of the utilized data. Validation is a context-dependent, multi-faceted challenge that is approached in various ways.
Experts were aware of some forms of variability unearthed in the PREP field study results, while also some contribution could be made. This is further described in sections 8.3 and 8.2.

8.2 Contribution of this Work

How are the ethnographic research results described in the last section to be understood and interpreted in respect to their possible uses in the design of EHR systems? The ability to formulate *implications for design* on the basis of ethnographic research, albeit commonly done, does not constitute a self-evident, imperative metric on the value of this research, as Paul Dourish remarks [79]. In addition to this practice, ethnographic research results can also unfold viability for the understanding of aspects of human computer interaction (HCI) in general as well as for designing technology specifically for a given setting *in themselves*; the analytic process and its reasoning and sense-making of data that originated from an ethnographic perspective can also hold insight, even if not directly translated into corresponding design guidelines or implications for design.
In this sense, the contribution of this thesis is meant to be understood two-fold:

A View on Practice

The field of EHR system usage in primary care is an intricately connected structure of technology usage and cooperative work that is subject to intensive research not only in light of the potential use of its data for purposes of secondary analysis. Semantic interoperability, as described by the interviewed experts, constitutes an ubiquitously important goal for next generation EHR systems, also because of "the increasing need for decision support to help clinicians cope with the overload of new medical evidence", "the importance of personalized healthcare and the need to engage patients more actively in the process of care" [22]. This requires a detailed understanding of the work processes, of technological and design aspects as well as, by any means not least, needs, skills experiences of the involved human beings.

As Hripcsak and Albers put it:

> The EHR is not a direct reflection of the patient and physiology, but a reflection of the recording process inherent in healthcare with noise and feedback loops. We must study the EHR as an object in itself, as if it were a natural system. This better understanding will then naturally support both broad-based outcome-oriented research and physiological research.[80, p. 119]

This relates to an interest in detailed views on EHR usage practices described in literature. In light of the observation of biasing effects QOF-related incentivization has on GP recording practices, Williams et al. call for a "*scientific research-based assessment of primary care data sources in general, specifically for the purpose of characterizing more accurately the strengths and weaknesses of the available data itself*" [61], a request seconded by Smith et al., who suggested qualitative work to investigate underlying reasons for variations in antibiotics prescribing practice among UK GPs found by them [52].

In order to transport the insight that was acquired in the course of the PREP field study's data analysis conducted in this master's thesis work in how GPs conduct medical consultations and include EHR systems in this highly professionalized work, this work exhibits an intentional verbosity and attention to detail, both to the end of documenting details of interest (e.g. the regional difference in RF blood tests described in section 7.4.2) and give a rich picture of the work practice.

Implications for Design

Apart from the view on practice, this master's thesis also offers implications for design as a result of the PREP field study analysis' reflection in light of aspects of secondary analysis. These design aspects do not reflect the results of this thesis as a whole, but represent interesting aspects that might translate into valuable design imperatives for the design of next generation EHR systems in respect of the requirements and needs of secondary analysts as well as the GPs who use them day by day. These implications are described in the next section.

8.3 Implications for Design

8.3.1 A Need for Interface and Interaction Guidelines

As described in section 7.4.5, a notable discrepancy emerged when reflecting the PREP field study's results of inter-system variability (described in chapter 5) against related literature and the interviewed experts' opinion: While differences in system design were acknowledged, little speculation would causally relate variability in EHR data to such differences, it would sometimes even be dismissed explicitly [61], [E5].

8.3.1.1 Selected, Exemplary EHR Interaction Design Differences

As described in chapter 5, differences in interface and interaction design of the three EHR systems at hand are various. Three examples are highlighted to illustrate this:

1) **Vision's code-centric design requires GPs to know a specific Read code vocabulary. SystmOne's code prediction feature is challenging for GPs bad at touch typing.**

GPs using SystmOne describe the activity of applying Read codes to a EHR documentation as cumbersome, not least because of the poorly designed interaction around SystmOne's code prediction feature. Among other aspects, it requires GPs to constantly observe the screen when entering text in order to dismiss possibly appearing Read code suggestions. GPs who describe themselves as bad at touch typing remarked, that they struggle with this, causing the annotation of EHR documentation with possibly unwanted, unintentional Read codes.

2) **The different approaches to interactionally scaffolding the usage of codes and free text cause variability in the amount and type of used codes.**

With Vision, it is obligatory to enter a Read code when documenting aspects of a medical consultation. This affects the amount of codes used, as illustrated in figure 8.1: GPs of surgery MS, using the EHR system Vision, use more codes than the other surgeries; also, the used codes are more diverse.

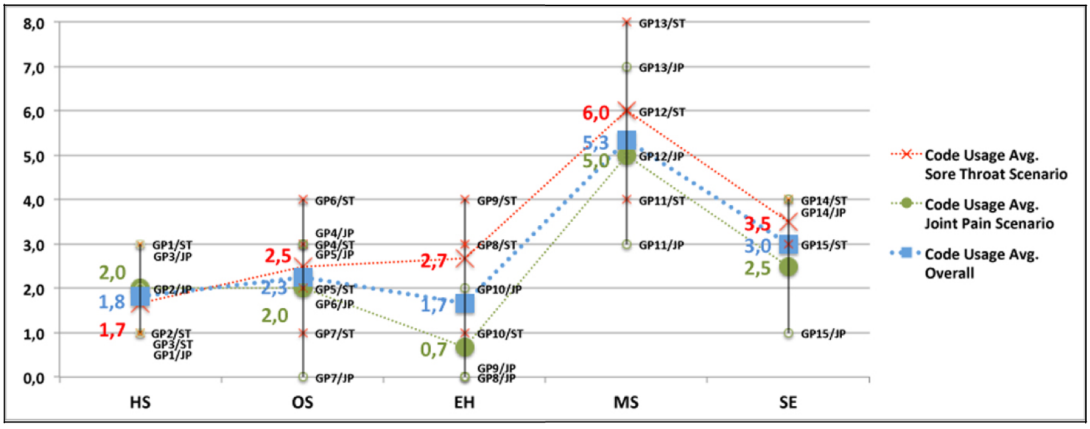


Figure 8.1: Surgery MS (Vision) exhibits higher code usage than other surgeries with other EHR systems. EMIS LV (Surgeries HS, EH and SE), SystmOne (Surgery OS) and Vision (Surgery MS).

3) **Interfaces infer a 'right' way to use them.**

As can be seen in figure 8.2, GPs from surgery OS, who are users of EHR system SystmOne, exhibit the least variability in their choice of how many categories they would use to document their consultations. Since the simulated patients invoked roughly the same medical responses of all participating GPs, making the documentation incentives of GPs from surgery OS roughly the same as with all other GPs, and since the usage of specific categories for documenting a consultation is not mandatory with any of the tested EHR systems, it stands to reason, that aspects peculiar to SystmOne incentivized GPs

of surgery OS to almost consistently use five categories to document their consultations, while GPs from other surgeries exhibited higher variability in this aspect.

Speculating on the reason for this, differences in interface design come to mind, as visible in figure 8.3: SystmOne prompts the user with four empty, white fields, which might imply to the user a necessity to 'fill them' with meaning. Vision on the other hand asks users for a Read code and subsequently recommends a possibly fitting category.

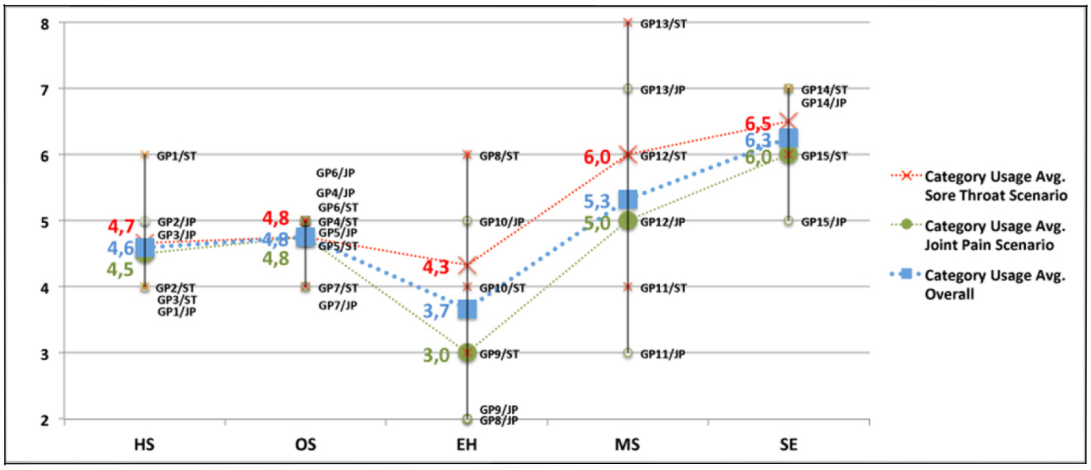


Figure 8.2: Overview of categorization of EHR documentation across surgeries.

8.3.1.2 Proposition of Guidelines

To mitigate against these forms of variability as well as instances of chance introduced to the usage practice of a system by poor design choices, the formulation of design guidelines might be a viable strategy. Such guidelines need to

1) **address aspects of interface and interaction design** around the usage of codes and categories for free text towards a unified concept of interaction. Putting aside at first the question of *which* approach to a category interface would yield data that more easily lends itself to purposes of secondary analysis - one like Vision, with more variability depending on user choices, or one like SystmOne, with little, expectable variability - it seems feasible, that standardization of (1) amount and (2) interactional affordance of categorization can improve data validity by reducing unnecessary variability. Similar thought applies to the interactional affordance of codification.

2) **be mindful of a pliable user experience**, e.g. in not demanding users who feel challenged by touch-typing to continuously look at the computer screen.

Further Remarks

It stands to reason, that the proposition of guidelines represent the right choice of mitigation against the variability across EHR systems found in the analysis of the

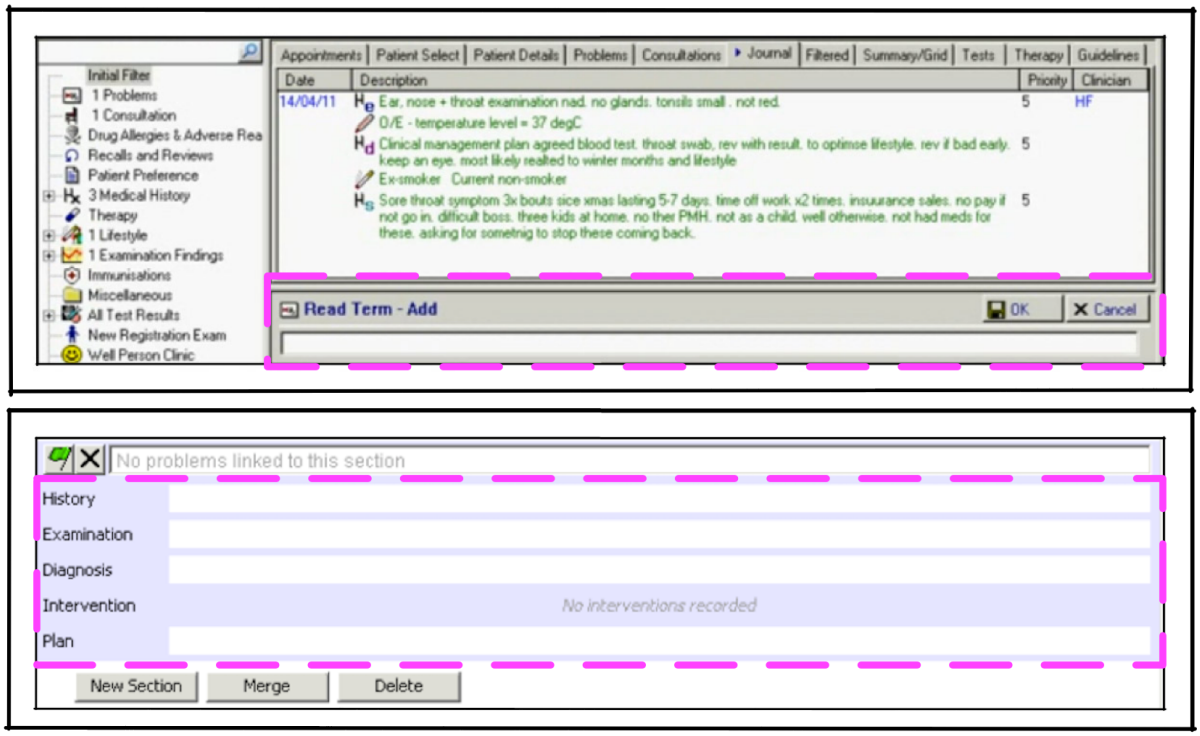


Figure 8.3: Comparing the interface of documentation structure with Vision (top) and SystmOne (bottom).

PREP field study. Harmonizing the usage of data structuring down to the interface and interaction design level is a sustainable choice also in light of general efforts of improving semantic interoperability across EHR systems. As Chan et al. remark in a conclusion on their literature review on EHR quality measures:

> "Poor data comparability will be an important impediment to the use of EHR data for quality reporting. To enhance data comparability, health care organizations will need to make careful decisions regarding the data source used for quality measurement, better standardize data documentation and extraction procedures, and implement effective interventions or tools to facilitate data harmonization across organizations."[11, p. 519]

Harmonization of interface and interaction of EHR systems furthermore might help understanding and reducing variability across EHR databases for purposes of secondary analysis. As described in section 7.4.4, variability across EHR databases is only partly explained; variability in EHR system interfaces and underlying interaction as found in this thesis might be involved in the causal background to such variability. Springate

et al. formulated a suspicion in this direction, but lacked empirical evidence to further explore EHR design as underlying reason for the variability in EHR data of databases CPRD - fed by EHR system Vision exclusively - and database DIN fed exclusively by EHR software Torex [77].
Not least, as remarked by the interviewed expert on NLP, can standardized categorization improve the utility of those categories for automated extraction tasks. The more consistently categories are used, the better they lend themselves as contextual meta-information in NLP extraction tasks, as described also by Porzel and Strube [81].

8.3.2 A Need for Disentangling Reimbursement Incentivization from Coding

Both from the expert interviews as well as the PREP field study's results it is evident, that the relevance of clinical routine data as primary care EHR documentation for external metrics, e.g. as a tool for surgeries for accessing reimbursement and/or as a tool for health care providers to monitor the establishment of various policies (like the Quality and Outcomes Framework (QOF)) exerts a biasing influence on the data. EHR system design needs to mitigate against this in order to improve its validity for secondary analysis purposes, ideally by completely disentangling this usage form from its clinical-medical purpose.

8.3.3 Secondary Analysis vs. Primary Use?

Another aspect that the PREP field study's analysis implies for the design of EHR systems is that, when designing such systems - an activity often driven by incentives of improving semantic interoperability for purposes of secondary analysis and transparency - the needs of the physicians who will end up using these systems needs to be kept in mind.

As Morrison et al. remark:

> There is good evidence for the secondary use benefits from structured and coded electronic health records (EHRs). The case has historically been presented that the use of structure and coding within EHRs also provides benefits to direct patient care, but the evidence base for this assertion has hitherto not been well documented. [22, p. 70]

As became evident in the PREP field study's analysis, structure in the sense of providing GPs with categories to document their consultations in is a valuable tool for them. The interviewed GPs also described the usage of codes as beneficial for purposes of usefulness like quickly searching through thick EHR documents. In respect to structure and coding within EHRs and its value for GPs, it seems important to mention, that usefulness for purposes of GP work is a multifaceted concept. While it may be useful to a surgery to use codes in order to document medical work for purposes of reimbursement, this

represents clerical work that ultimately accumulates to a substantial amount of the time GPs can devote to each individual patient. Thus, EHR systems need to be designed for a quick, pliable use and impose only minimal cognitive load on the user. This might seem trivial as an implication, but needs to be stressed nonetheless as an important imperative in relation to design work being incentivized by goals that are formulated and evaluated on entirely different scales, e.g. data usability for purposes of secondary analysis.

8.4 Limitations

The results of this work exhibit limitations that need to be kept in mind when interpreting its conclusions: (1) Qualitative, ethnographical insight into a subject always exhibits an element of subjectivity; the analysis conducted on and the results drawn from the PREP field study's data, while conducted carefully and diligently, is interpretative of the observed phenomena and thus can not be considered the only possible view on the material at hand. (2) The comparative interpretation of usage practice of EHR systems conducted in chapter 5 constitutes empirical evidence of variability, which hints on differences across systems and surgeries, but does not constitute proof of them, as for this, more data would be required, especially with EHR systems Vision and SystmOne, for which only one surgery each was available for analysis in this master's thesis.

8.5 Conclusion and Future Work

This master's thesis investigated the usage practice of EHR systems in primary care medical consultations and reflected on them from the perspective of secondary analysis, contributing insight in the usage practice around EHR systems to the imperative that EHR systems need to be designed with appreciation of this practice and the needs of the people using them; this to the end of creating sustainable technology that serves their needs and fulfills its purpose also in light of secondary data usage.

Returning to criteria for evaluating qualitative work postulated by Charmaz described in section 3.1 [27], it is hoped, that the detailed view on the GPs work with the simulated patients and all related data has created a *credible* and in part even *original* (that is, unknown to the state-of-the-art in the scientific community) work that will prove to be *useful* for EHR design.

Future Work

In the course of this master's thesis, many aspects of HCI in EHR systems were touched that might inspire future work. The analysis conducted on the PREP field study's data material indicated starting points that were not explored in this master's thesis, but could pose an interesting source for further insight: A language analysis on the EHR documentation, especially its free text portions, might hold further insight in usage practice of GPs when working with EHR systems. Further investigation of the pattern of incompleteness in rheumatoid factor (RF) blood tests might hold important implications

on what to make of this indicator (or its absence) in secondary analysis, as it is considered Missing Not at Random (MNAR).

Bibliography

[1] D. Swinglehurst, T. Greenhalgh, and C. Roberts, "Computer Templates in Chronic Disease Management: Ethnographic Case Study in General Practice," *BMJ Open*, vol. 2, no. 6, 2012. Available online: `http://bmjopen.bmj.com/content/2/6/e001754.full.pdf+html`, accessed 15-November-2015.

[2] C. Heath and P. Luff, "Documents and Professional Practice: 'Bad' Organisational Reasons for 'Good' Clinical Records," *Computer Supported Cooperative Work*, pp. 354–363, 1996.

[3] D. Swinglehurst, T. Greenhalgh, M. Myall, and J. Russell, "Ethnographic Study Of ICT-Supported Collaborative Work Routines in General Practice," *BMC Health Services Research*, vol. 10, no. 1, p. 348, 2010. Available online: `http://www.biomedcentral.com/1472-6963/10/348`, accessed 15-November-2015.

[4] G. Hardstone, M. Hartswood, R. Procter, R. Slack, A. Voss, and G. Rees, "Supporting Informality: Team Working and Integrated Care Records," in *CSCW '04: Proceedings of the 2004 ACM conference on Computer supported cooperative work*, (New York, New York, USA), p. 142, ACM Request Permissions, 2004.

[5] T. Greenhalgh, H. Potts, G. Wong, P. Bark, and D. Swinglehurst, "Tensions and Paradoxes in Electronic Patient Record Research: A Systematic Literature Review Using the Meta-narrative Method," *Milbank Quarterly*, vol. 87, no. 4, pp. 729–788, 2009.

[6] M. S. Feldman and B. T. Pentland, "Reconceptualizing Organizational Routines as a Source of Flexibility and Change," *Administrative Science Quarterly*, vol. 48, no. 1, pp. 94–118, 2003. Available online: `http://asq.sagepub.com/content/48/1/94.abstract`, accessed 15-November-2015.

[7] C. Roberts and S. Sarangi, "Theme-oriented discourse analysis of medical encounters," *Medical Education*, vol. 39, no. 6, pp. 632–640, 2005.

[8] S. Boslaugh, "An Introduction to Secondary Data Analysis," *Secondary Data Sources for Public Health: A Practical Guide. Cambridge University Press.*, 2007.

[9] N. F. Khan, S. E. Harrison, and P. W. Rose, "Validity of Diagnostic Coding Within the General Practice Research Database: A Systematic Review.," *British Journal of General Practice*, vol. 60, no. 572, pp. e128–36, 2010.

[10] N. G. Weiskopf and C. Weng, "Methods and Dimensions of Electronic Health Record Data Quality Assessment: Enabling Reuse for Clinical Research.," *Journal of American Medical Informatics Association*, vol. 20, no. 1, pp. 144–151, 2013.

[11] K. S. Chan, J. B. Fowles, and J. P. Weiner, "Review: Electronic Health Records and the Reliability and Validity of Quality Measures: a Review of the Literature," *Medical care research and review : MCRR*, vol. 67, no. 5, pp. 503–527, 2010.

[12] T. Botsis, G. Hartvigsen, F. Chen, and C. Weng, "Secondary Use of EHR: Data Quality Issues and Informatics Opportunities," *Summit on Translational Bioinformatics. 2010*, pp. 1–5, 2010.

[13] N. G. Weiskopf, G. Hripcsak, S. Swaminathan, and C. Weng, "Defining and Measuring Completeness of Electronic Health Records for Secondary Use," *Journal of Biomedical Informatics*, vol. 46, no. 5, pp. 830–836, 2013.

[14] A. R. Tate, A. G. R. Martin, A. Ali, and J. A. Cassell, "Using Free Text Information to Explore How and When GPs Code a Diagnosis of Ovarian Cancer: An Observational Study Using Primary Care Records of Patients With Ovarian Cancer," *BMJ Open*, vol. 1, no. 1, pp. e000025–e000025, 2011.

[15] A. Nicholson, E. Ford, K. A. Davies, H. E. Smith, G. Rait, A. R. Tate, I. Petersen, and J. Cassell, "Optimising Use of Electronic Health Records to Describe the Presentation of Rheumatoid Arthritis in Primary Care: A Strategy for Developing Code Lists," *Plos One*, vol. 8, no. 2, 2013.

[16] L. Axelrod, G. Fitzpatrick, F. Henwood, J. Cassell, H. Smith, A. Nicholson, and G. Rait, "Data Recording in Primary Care Field Studies: Patient Records Enhancement Project," in *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2011 5th International Conference on*, pp. 563–567, May 2011.

[17] L. Axelrod, G. Fitzpatrick, F. Henwood, L. Thackray, B. Simpson, A. Nicholson, H. Smith, G. Rait, and J. Cassell, "'Acted Reality' in Electronic Patient Record Research: A Bridge Between Laboratory and Ethnographic Studies," in *Proceedings of the 13th IFIP TC 13 International Conference on Human-computer Interaction - Volume Part II*, INTERACT'11, pp. 73–80, Berlin, Heidelberg: Springer-Verlag, 2011.

[18] S. Gold, *Electronic health records: the main patient record providers*. The Guardian, 2011. available online: `http://www.theguardian.com/healthcare-network/2011/apr/27/electronic-health-records-cerner-isoft-emis`, accessed 15-December-2015.

[19] S. O'Hanlon, *EMIS Web - most widely used GP system in England.* Patient.info, 2013. available online: http://patient.info/blogs/pro/2013/03/emis-web---most-widely-used-gp-system-in-england, accessed 15-December-2015.

[20] T. Meek, *Record sharing: TPP and Emis integrate.* digitalhealth.net, 2015. available online: http://www.digitalhealth.net/shared_care_records/46802/gp-record-sharing:-suppliers-integrate, accessed 15-December-2015.

[21] H. Bosse, M. Nickel, S. Huwendiek, J. Junger, J. Schultz, and C. Nikendei, "Peer Role-Play and Standardised Patients in Communication Training: A Comparative Study On the Student Perspective on Acceptability, Realism and Perceived Effect," *BMC Medical Education*, vol. 10, no. 1, p. 27, 2010.

[22] Z. Morrison, B. Fernando, D. Kalra, K. Cresswell, A. Robertson, A. Hemmi, and A. Sheikh, *An Evaluation of Different Levels of Structuring Within the Clinical Record. Final report for the NHS Connecting for Health Evaluation Programme.* NHS Connecting for Health Evaluation Programme (NHS CFHEP), 2012.

[23] S. de Lusignan, "Codes, Classifications, Terminologies and Nomenclatures: Definition, Development and Application in Practice," *Informatics in Primary Care*, pp. 65–70, 2005.

[24] E. Guba and Y. Lincoln, "Competing Paradigms in Qualitative Research," in *Handbook of Qualitative Research* (N. Denzin and Y. Lincoln, eds.), pp. 105–117, Sage, Thousand Oaks, 1994.

[25] V. Braun and V. Clarke, "Using Thematic Analysis in Psychology," *Qualitative Research in Psychology*, vol. 3, no. 2, pp. 77–101, 2006. Available online: http://www.tandfonline.com/doi/abs/10.1191/1478088706qp063oa, accessed 15-November-2015.

[26] P. Krüger and I. Meyer, "Review Essay: A Journey through Grounded Theory," *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, vol. 8, no. 1, 2007.

[27] K. Charmaz, *Constructing Grounded Theory*, vol. 2. SAGE Publications Ltd, 2006.

[28] H. Knoblauch, A. Baer, E. Laurier, S. Petschke, and B. Schnettler, "Visual analysis. new developments in the interpretative analysis of video and photography," *Forum: Qualitative Social Research*, vol. 9, no. 3, 2008.

[29] K. Eder, "Videographie im Forschungsprojekt die Wiener Kinderkrippenstudie," *University of Vienna, diploma-thesis in educational sciences*, 2011. Available online at: http://othes.univie.ac.at/15089/1/2011-04-21_8606036.pdf; accessed 15-November-2015.

[30] H. Knoblauch, "Videography. Focused Ethnography and Video Analysis," *Video Analysis: Methodology and Methods. Qualitative Audiovisual Data Analysis in Sociology*, pp. 69–83, 2006. Quoted after [29].

[31] C. Jewitt, "An Introduction to Using Video for Research," *MODE node, Institute of Education, London, National Centre of Research Methods (NCRM) Working paper*, 2012.

[32] R. Goldman, "Video Representations and the Perspectivity Framework: Epistemology, Ethnography, Evaluation and Ethics," in *Video Research in the Learning Sciences*, Taylor & Francis, 2014.

[33] A. Garcez, R. Duarte, and Z. Eisenberg, "Production and analysis of video recordings in qualitative research," *Educacao e Pesquisa*, vol. 37, pp. 249 – 261, 2011.

[34] H. Lomax and N. Casey, "Recording Social Life: Reflexivity and Video Methodology," *Sociological Research Online*, vol. 3, no. 2, 1998. Available online at: `http://www.socresonline.org.uk/3/2/1.html`; accessed 15-November-2015.

[35] K. Mruck and F. Breuer, "Subjectivity and Reflexivity in Qualitative Research," *Forum Qualitative Social Research*, vol. 4, no. 2, 2003. Available online at: `http://www.qualitative-research.net/index.php/fqs/article/view/696/1504`; accessed 15-November-2015.

[36] H. Knoblauch, "Focused ethnography," *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, vol. 6, no. 3, 2005. Available online at: `http://www.qualitative-research.net/index.php/fqs/article/view/20`; accessed 15-November-2015.

[37] A. M. F. d. A. Sadalla and P. Larocca, "Autoscopia: um procedimento de pesquisa e de formacao," *Educacao e Pesquisa*, vol. 30, pp. 419 – 433, 12 2004. Quoted after [33].

[38] M. Wagner-Willi, "Die Dokumentarische Videointerpretation in der erziehungswissenschaftlichen Ethnographieforschung," *Ethnographie und Erziehungswissenschaft. Methodologische Reflexionen und empirische Annäherungen*, pp. 221–231, 2009. Quoted after [29].

[39] H. Knoblauch, "Die Video-Interaktions-Analyse," *Sozialer Sinn*, vol. 5, no. 1, pp. 123–138, 2004. Available online at: `http://nbn-resolving.de/urn:nbn:de:0168-ssoar-7571`, accessed 15-November-2015.

[40] J. Dinkelaker and M. Herrle, "Erziehungswissenschaftliche Videographie. Eine Einführung," *Verlag für Sozialwissenschaften: Wiesbaden*, 2009. Quoted after [29].

[41] N. Denzin, "Symbolischer Interaktionismus," *Qualitative Forschung. Ein Handbuch*, 2007.

[42] U. Flick, *An Introduction to Qualitative Research*. Sage Publications Ltd., fourth ed., 2009.

[43] B. Cheek, *Calgary-Cambridge guide to the medical interview - communication process*. 2011. available online: `http://www.gp-training.net/training/communication_skills/calgary/guide.htm`, accessed 15-December-2015.

[44] S. Kurtz, J. Silverman, and J. Draper, *Teaching and Learning Communication Skills in Medicine*. Radcliffe Medical Press (Oxford), 1998. quoted after [43].

[45] *General practice in the UK: Media Brief*. British Medical Association (BMA), 2014. available online: `http://is.gd/SM5SBG`, accessed 15-January-2015.

[46] E. Herrett, S. L. Thomas, W. M. Schoonen, L. Smeeth, and A. J. Hall, "Validation and Validity of Diagnoses in the General Practice Research Database: A Systematic Review," *British Journal of Clinical Pharmacology*, vol. 69, no. 1, pp. 4–14, 2010.

[47] K. Häyrinen, K. Saranto, and P. Nykänen, "Definition, Structure, Content, Use and Impacts of Electronic Health Records: A Review Of the Research Literature," *International Journal of Medical Informatics*, vol. 77, no. 5, pp. 291–304, 2008.

[48] J. Heaton, "Secondary analysis of qualitative data," *Social Research Update*, vol. 22, 1998.

[49] S. E. Gabriel and K. Michaud, "Epidemiological Studies in Incidence, Prevalence, Mortality, and Comorbidity of the Rheumatic Diseases," *Arthritis Research & Therapy*, vol. 11, no. 3, pp. 1–16, 2009.

[50] K. Sheikh, L. Gilson, I. A. Agyepong, K. Hanson, F. Ssengooba, and S. Bennett, "Building the Field of Health Policy and Systems Research: Framing the Questions," *PLoS Med*, vol. 8, p. e1001073, 08 2011.

[51] M. Lawrence, A. Coulter, and L. Jones, "A total audit of preventive procedures in 45 practices caring for 430,000 patients.," *BMJ*, vol. 300, no. 6738, pp. 1501–1503, 1990.

[52] S. Smith, G. Smith, H. Heatlie, J. Bashford, D. Ashcroft, N. Verlander, G. Duckworth, B. Mason, B. Smyth, and S. Maxwell, "Reducing variation in antibacterial prescribing rates for 'cough/cold' and sore throat between 1993 and 2001: Regional analyses using the general practice research database," *Public Health*, vol. 120, no. 8, pp. 752 – 759, 2006.

[53] N. Dunn, D. Lane, H. Everitt, and P. Little, "Use of antibiotics for sore throat and incidence of quinsy," *British Journal of General Practice*, vol. 57, no. 534, pp. 45–49, 2007.

[54] C. J. Edwards, J. Campbell, T. Van Staa, and N. K. Arden, "Regional and Temporal Variation in the Treatment of Rheumatoid Arthritis Across the UK: A Descriptive Register-Based Cohort Study," *BMJ Open*, vol. 2, no. 6, 2012.

[55] A. Judge, G. Wallace, D. Prieto-Alhambra, N. K. Arden, and C. J. Edwards, "Can the Publication of Guidelines Change the Management of Early Rheumatoid Arthritis? An Interrupted Time Series Analysis From the United Kingdom," *Rheumatology (Oxford, England)*, p. kev268, 2015.

[56] A. S. Forster, H. Dodhia, H. Booth, A. Dregan, F. Fuller, J. Miller, C. Burgess, L. McDermott, and M. C. Gulliford, "Estimating the yield of NHS Health Checks in England: a population-based cohort study," *Journal of public health (Oxford, England)*, vol. 37, no. 2, pp. 234–240, 2015.

[57] M. Tolar and E. Balka, "Caring for individual patients and beyond: Enhancing care through secondary use of data in a general practice setting," *International Journal of Medical Informatics*, vol. 81, no. 7, pp. 461–474, 2012.

[58] E. Ford, A. Nicholson, R. Koeling, A. R. Tate, J. Carroll, L. Axelrod, H. E. Smith, G. Rait, K. A. Davies, I. Petersen, T. Williams, and J. A. Cassell, "Optimising the Use of Electronic Health Records To Estimate The Incidence of Rheumatoid Arthritis in Primary Care: What Information is Hidden in Free Text?," *BMC Medical Research Methodology*, vol. 13, p. 1, Aug. 2013.

[59] A. R. Tate, N. Beloff, B. Al-Radwan, J. Wickson, S. Puri, T. Williams, T. Van Staa, and A. Bleach, "Exploiting the Potential Of Large Databases of Electronic Health Records For Research Using Rapid Search Algorithms and an Intuitive Query Interface.," *Journal of the American Medical Informatics Association*, vol. 21, no. 2, pp. 292–298, 2014.

[60] I. M. Carey, D. G. Cook, S. De Wilde, S. A. Bremner, N. Richards, S. Caine, D. P. Strachan, and S. R. Hilton, "Developing a Large Electronic Primary Care Database (Doctors' Independent Network) For Research.," *International Journal of Medical Informatics*, vol. 73, no. 5, pp. 443–453, 2004.

[61] T. Williams, T. Van Staa, S. Puri, and S. Eaton, "Recent Advances in the Utility and Use of the General Practice Research Database As an Example of a UK Primary Care Data Resource," *Therapeutic advances in drug safety*, vol. 3, no. 2, pp. 89–99, 2012.

[62] E. Herrett, A. D. Shah, R. Boggon, S. Denaxas, L. Smeeth, T. Van Staa, A. Timmis, and H. Hemingway, "Completeness and diagnostic validity of recording acute myocardial infarction events in primary care, hospital care, disease registry, and national mortality records: cohort study," *Bmj-British Medical Journal*, vol. 346, no. may20 3, 2013.

[63] S. Thomas, C. Edwards, L. Smeeth, C. Cooper, and A. Hall, "How accurate are diagnoses for rheumatoid arthritis and juvenile idiopathic arthritis in the General Practice Research Database?," vol. 46, Oxford University Press (OUP), 2007.

[64] K. I. Morley, J. Wallace, S. C. Denaxas, R. J. Hunter, R. S. Patel, P. Perel, A. D. Shah, A. D. Timmis, R. J. Schilling, and H. Hemingway, "Defining Disease Phenotypes Using National Linked Electronic Health Records: A Case Study of Atrial Fibrillation," *Plos One*, vol. 9, no. 11, 2014.

[65] R. Pivovarov and N. Elhadad, "Automated Methods for the Summarization of Electronic Health Records," *Journal of the American Medical Informatics Association*, vol. 22, pp. ocv032–947, Apr. 2015.

[66] J.-H. Lin and P. J. Haug, "Exploiting Missing Clinical Data in Bayesian Network Modeling For Predicting Medical Problems," *Journal of Biomedical Informatics*, vol. 41, no. 1, pp. 1–14, 2008.

[67] C. J. Sammon, A. Miller, K. R. Mahtani, T. A. Holt, N. J. McHugh, R. A. Luqmani, and A. L. Nightingale, "Missing Laboratory Test Data in Electronic General Practice Records: Analysis of Rheumatoid Factor Recording in the Clinical Practice Research Datalink," *Pharmacoepidemiology and Drug Safety*, vol. 24, no. 5, pp. 504–509, 2015.

[68] E. P. Vamos, U. J. Pape, A. Bottle, F. L. Hamilton, V. Curcin, A. Ng, M. Molokhia, J. Car, A. Majeed, and C. Millett, "Association of Practice Size and Pay-For-Performance Incentives With the Quality of Diabetes Management in Primary Care," *Canadian Medical Association Journal*, vol. 183, no. 12, pp. 809–816, 2011.

[69] C. A. Chew-Graham, C. Hunter, S. Langer, A. Stenhoff, J. Drinkwater, E. A. Guthrie, and P. Salmon, "How QOF is Shaping Primary Care Review Consultations: A Longitudinal Qualitative Study," *BMC Family Practice*, vol. 14, no. 1, pp. 1–1, 2013.

[70] R. Mathur, K. Bhaskaran, N. Chaturvedi, D. A. Leon, T. vanStaa, E. Grundy, and L. Smeeth, "Completeness and Usability of Ethnicity Data in UK-based Primary Care and Hospital Databases," *Journal of public health (Oxford, England)*, vol. 36, no. 4, pp. 684–692, 2014.

[71] T. A. Simon, A. L. Smitten, J. Franklin, J. Askling, D. Lacaille, F. Wolfe, M. C. Hochberg, K. Qi, and S. Suissa, "Malignancies in the rheumatoid arthritis abatacept clinical development programme: an epidemiological assessment," *Annals of the Rheumatic Diseases*, vol. 68, no. 12, pp. 1819–1826, 2009.

[72] H. W. Frank H. Mader, *Der Assistenzarzt in der Allgemeinpraxis*. Springer Berlin (Deutschland), 1998.

[73] W. Fink and G. Haidinger, "Die Häufigkeit von Gesundheitsstörungen in 10 Jahren Allgemeinpraxis," *ZFA - Zeitschrift für Allgemeinmedizin*, vol. 83, pp. 102–108, 2007.

[74] R. Koeling, A. R. Tate, and J. A. Carroll, "Automatically Estimating the Incidence of Symptoms Recorded in GP Free Text Notes," in *Proceedings of the First International Workshop on Managing Interoperability and Complexity in Health Systems*, MIXHS '11, (New York, NY, USA), pp. 43–50, ACM, 2011.

[75] R. Koeling, J. Carroll, R. Tate, and A. Nicholson, "Annotating a Corpus of Clinical Text Records for Learning to Recognize Symptoms Automatically," in *Proceedings of LOUHI 2011 Third International Workshop on Health Document Text Mining and Information Analysis*, vol. 744 of *CEUR Workshop Proceedings*, pp. 43–50, Trondheim, Norway: Norwegian University of Science and Technology, August 2011.

[76] P. R. Rijnbeek, "Converting To a Common Data Model: What is Lost in Translation?," *Drug Safety*, vol. 37, no. 11, pp. 893–896, 2014.

[77] D. A. Springate, D. M. Ashcroft, E. Kontopantelis, T. Doran, R. Ryan, and D. Reeves, "Can Analyses of Electronic Patient Records be Independently and Externally Validated? Study 2 - The Effect of Beta-Adrenoceptor Blocker Therapy On Cancer Survival: A Retrospective Cohort Study," *BMJ Open*, vol. 5, no. 4, 2015.

[78] E. S. Petherick, K. E. Pickett, and N. A. Cullum, "Can different primary care databases produce comparable estimates of burden of disease: results of a study exploring venous leg ulceration," *Family Practice*, vol. 32, pp. 374–380, Aug. 2015.

[79] P. Dourish, "Implications for Design," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '06, (New York, NY, USA), pp. 541–550, ACM, 2006.

[80] G. Hripcsak and D. J. Albers, "Next-generation phenotyping of electronic health records.," *Journal of the American Medical Informatics Association*, vol. 20, no. 1, pp. 117–121, 2013.

[81] M. S. Robert Porzel, "Towards Context-adaptive Natural Language Processing Systems," in *Computational Linguistics for the New Millennium: Divergence or Synergy?* (H. V. Manfred Klenner, ed.), Peter Lang Intl. Publishers, 2002.

APPENDIX A

Appendix: Video Analysis Process

A.1 Field Study Data Sources

Table A.1 contains a mapping of the GP labels used in this work and their counterpart in the lists and folder structure of the PREP field study's data material.

A.2 Example EHR Documentation of *Scott* and *Jane* case

Underlined portions denote Read-coded information.

A.2.1 Joint Pain Scenario

EMIS LV:

GP1
History: Few months deteriorating joints hands and feet symmetrically, aching pain and stiffness, worse in morning and better through day, accountant-use of computer and worried about this, tried wrist splinting no effect, also ibu sporadically no effect gran ?RA, otherwise well though feels run down with this
Examination: Swollen DIPs and PIPs both hands, tender to touch, no other ex possible
Comment: Exclude AI cause, bloods initially esr/crp/fbc (not RA as per rheum wishes), trial glucosamine and back for blood results
Problem Title: Hand joint stiffness (minor)

SystmOne:

GP4
History: bilateral stiff, swollen hands and feet / ankles for some time. accountant and impacting on ability to type. morning stiffness. has tried ibuprofen – not much help.
Examination: MCPJ swelling, warm. no deformity. stiff hands
Medication: Naproxen 500 mg tablets – 28 tablets – take one twice daily
Pathology Request: (Blood test): CRP Free T4 Random Glucose TSH Electrolytes ESR Full Blood Count
Plan: consider rheumatoid – check bloods I will contact to discuss. given naproxen. mentioned GI effects. Arthralgia NOS

Nr.	ID / System / Scenario	Video Folder/File	ID in [EHRTr]	WT-File	SR-File
1	GP1 / EMIS / JP	hs sp gp 10	HS 10 EMIS LV	HS01_EMIS_Walkthrough	HS10 screen
2	GP1 / EMIS / ST	hs sp gp 10	HS 10 EMIS LV	HS01_EMIS_Walkthrough	HS10 screen
3	GP2 / EMIS / JP	hs sp gp 09	HS 09 EMIS LV	*included in video file*	hs09 screen capture
4	GP2 / EMIS / ST	hs sp gp 09	HS 09 EMIS LV	*included in video file*	hs09 screen capture
5	GP3 / EMIS / JP	hs 08 original recordings	HS 08 EMIS LV	*included in video file*	HS08REG screen 2
6	GP3 / EMIS / ST	hs 08 original recordings	HS 08 EMIS LV	*included in video file*	HS08REG screen 2
7	GP4 / SystemOne / JP	os/gpw 01 so 27 06 11	OS 01/21 ew	*included in video file*	*not available*
8	GP4 / SystemOne / ST	os/gpw 01 so 27 06 11	OS 01/21 ew	*included in video file*	*not available*
9	GP5 / SystemOne / JP	os/gpr 02 so 27 06 11	OS 01/22 fr	*included in video file*	*not available*
10	GP5 / SystemOne / ST	os/gpr 02 so 27 06 11	OS 01/22 fr	*included in video file*	*not available*
11	GP6 / SystemOne / JP	os/os gp03 27 06 11	OS 01/24 pl	*included in video file*	*not available*
12	GP6 / SystemOne / ST	os/os gp03 27 06 11	OS 01/24 pl	*included in video file*	*not available*
13	GP7 / SystemOne / JP	os/os gp04 27 06 11	OS 01/23 sm	*included in video file*	*not available*
14	GP7 / SystemOne / ST	os/os gp04 27 06 11	OS 01/23 sm	*included in video file*	*not available*
15	GP8 / EMIS / JP	eh02_EMIS	EH 02 EMIS LV	*included in video file*	eh02_screen
16	GP8 / EMIS / ST	eh02_EMIS	EH 02 EMIS LV	*included in video file*	eh02_screen
17	GP9 / EMIS / JP	eh04_EMIS	EH 04 EMIS LV	*included in video file*	eh04_screen
18	GP9 / EMIS / ST	eh04_EMIS	EH 04 EMIS LV	*included in video file*	eh04_screen
19	GP10 / EMIS / JP	eh05_EMIS	EH 05 EMIS LV	*included in video file*	MM screen capture
20	GP10 / EMIS / ST	eh05_EMIS	EH 05 EMIS LV	*included in video file*	MM screen capture
21	GP11 / VISION / JP	ms_01_VISION	MS Vis/18et	*included in video file*	ms01_screen
22	GP11 / VISION / ST	ms_01_VISION	MS Vis/18et	*included in video file*	ms01_screen
23	GP12 / VISION / JP	ms_02_VISION	MS Vis/20 *(sic)*	*included in video file*	ms02_screen
24	GP12 / VISION / ST	ms_02_VISION	MS Vis/17hf *(sic)*	*included in video file*	ms02_screen
25	GP13 / VISION / JP	ms_04_VISION	MS Vis/17hf *(sic)*	*included in video file*	*not available*
26	GP13 / VISION / ST	ms_04_VISION	MS Vis/20 *(sic)*	*included in video file*	*not available*
27	GP14 / EMIS / JP	Se_08_EMIS	SE 08 EMIS LV	*included in video file*	Se08_screen
28	GP14 / EMIS / ST	Se_08_EMIS	SE 08 EMIS LV	*included in video file*	Se08_screen
29	GP15 / EMIS / JP	Se_09_EMIS	SE 09 EMIS LV	*included in video file*	se09_screen
30	GP15 / EMIS / ST	Se_09_EMIS	SE 09 EMIS LV	*included in video file*	se09_screen

Table A.1: List of sources to the PREP field study's data used in this work including walkthrough and screen recording files.

Vision:

GP11
Intervention/Tablet-Icon: NAPROXEN tabs 500mg Supply (28) tablet(s) 1 TWICE A DAY AFTER FOOD
History/Diagnosis-Icon: Pain in joint – arthralgia 3-4 weeks pain in multiple joints of hands and feet. generally feeling tired and run down. No fever. No h/o foreign travel or viral/gastro illness. Work as accountant. Taking ibuprofen with little benefit. *[lb]* FH arthritis – grandmother ?what typre – was very disabled by it *[lb]* OE swollen fingers, swelling of MTP and PIP joints *[lb]* Bt and early rv
Vial-Icon: Blood test due Joint pains Rheumatoid factor, FBC, U+E, LFT, ESR, CRP

A.2.2 Sore Throat Scenario

EMIS LV:

GP2
Problem Title: Pharyngitis Recurrent
History: 2 yrs every 3-4 months. 3-4 days off work because of sore throat following on a coryzal illness. Concerned as an insurance salesman commission based. Twins at 3 and 5 years old at school.
Examination: No nodes. Injected pharynx. Tonsils NAD
Comment: Swab Blood tests. Discussion general self care and stress management

---

SystmOne:

GP5
History: Sore throat on and offfor the past three years. Three episodes per year. Starting to impact on work. Boss not very happy. He has three young children. He is concerned that his immune system may be low. No night sweats. no wt loss. Well in between sore throats.
Examination: INflammed throat, no particular swelling.
Medication: Benzydamine 0.15% mouthwash sugar free – 300 millilitres – 15ml every 3 hrs
Pathology Request: (Blood test:) Full Blood Count, Glandular fever Screen
Plan: Difflam for throat. Blood tests to test for glandular fever and FBC. Rv with results.

Vision:

GP12
History/Symptom-Icon: Sore throat symptom 3x bouts since xmas lasting 5-7 days. time off work. X2 times. Insurance sales. No pay if not go in. difficult boss. Three kids at home. no ther PMH . not as a child. Well otherwise. Not had meds for these. Asking for sometnig to stop these coming back
Cigarette-Icon: Ex smoker Current non smoker
History/Diagnosis-Icon: Clinical management plan agreed blood test. throat swab, rev with result. to optimize lifestyle. rev if bad early. Keep an eye. Most likely related to winter months and lifestyle
Pencil-Icon: O/E – temperature level = 37 degC
History/Examination-Icon: Ear nose + throat examination nad. no glands. tonsils small. not red.
Vial-Icon: Blood test due (recurrent sore throat) fbc, throat swab

---

APPENDIX B

Appendix: Expert Interviews

B.1 Interview Outline - Epidemiologist

1. Introduction
2. General Information / any Questions beforehand re. the info in the E-Mail / Consent-Form
3. What is your research interest in epidemiology?
4. Which data do you use as basis for your work?
5. What quality demands do you have on that data / How do you check the quality of the data? / What things do you look for to have confidence in the data?
6. What strengths and weaknesses of primary care (pc) data do you find most important?
7. Show sheet with stills of all 15 examined recordings as a visual anchor, explain setup of simulated patients and two scenarios. Give resulted, condensed EHR-data to skim through.
8. How diverse would you expect the results?
9. Could you please look through those and tell me if anything / Do any differences look striking to you?
10. Is this diversity in ehr documentation for the „same“ patient for you expectable?
11. Mention reasons for differences according to video analysis. Core reasons anyway, more specific reasons depending on research interest / in context with answers.

Core reasons: Interpretation and construction in consultation, different EHR-systems yield different documentation, chance More specific e.g. reasons for free text-diversity, ..

12. How relevant are these circumstances for your work? E.g. in terms of diagnosis validation?

13. What would be examples of secondary analyses where such differences and reasons would be relevant?

14. Any other issues or thoughts that come to mind about the data they use? Thank you very much for your time.

15. Offer to keep in the loop re results.

B.2 Interview Transcript Example - Epidemiologist

Expert Interview with E1.

For an overview of experts, see section 3.5, e.g. figure 3.6.

Meine erste Frage wäre: Welche Daten verwenden Sie für Ihre Epidemiologischen Forschungen?

Da gibt's prinzipiell zwei verschiedene Datenquellen. Das eine sind offizielle Daten der Republik Österreich über Statistik Austria, da verwenden wir vor allem Mortalitätsdaten, aber auch Krebsregisterdaten, und Morbiditätsdaten, Krankenhausstatistiken zum Beispiel.

Ok.

..und das andere sind Daten, die wir selber erheben: Einerseits im Feld, mit großen Studien, also z.B. die ISAAC-Studie, das war die International Study on Asthma and Allergy in Childhood, wo die Häufigkeit von Asthma und Allergien bei Kindern erhoben wurde. Also, selbst direkt Daten erheben, und andererseits haben wir beispielsweise Studien gemacht in einem Krankenhaus in Amstetten, wo Personen befragt wurden, was die Beweggründe sind, die Ambulanzen aufzusuchen. Und im Zusammenhang damit, aber auch unabhängig davon kommt's immer wieder dazu, dass Allgemeinmediziner kommen und sagen, "Wir haben so tolle Daten, wollen wir das nicht auswerten".

Alles klar. Und wenn Sie Daten verwenden, die aus einem Computersystem kommen, einem System, das von den Ärzten selbst verwenden worden ist, um z.B. Diagnosen zu dokumentieren und Behandlungsverläufe, welche Stärken und Schwächen sehen Sie in diesen Daten?

Da kommt's sehr drauf an. Ich habe vor einigen Jahren versucht, ein Morbiditätsregister aufzuziehen. Geld von der Nationalbank bekommen, 2x. System auf die Beine gestellt, das automatisiert aus der EDV der Ärzte, Daten absaugt, diese Datensätze aufteilt in einerseits

personalisierte Daten, die an Statistik Austria gingen, die haben das Anonymisiert und wieder an uns geleitet, und die Behandlungsdaten wurden direkt an uns gesendet. Diese Daten sind.. denen vertraue ich blind. Weil ich mit den Ärzten genau besprochen habe, was sie eintragen, und wie sie's eintragen. Ich weiß aber auch, dass Daten, die routinemäßig von Sozialversicherungsträgern usw. erfasst werden, dass die nicht die wahre Situation wiedergeben. Es wurde extra bei diesem Projekt von den Ärzten angemerkt, dass wir nicht die offiziellen Diagnosedaten verwenden sollen, sondern das, was die Ärzte für sich in ihrer Kartei eintragen.

Interessant. Wie wurde das begründet?

Das steht damit im Zusammenhang, dass an Sozialversicherungsträger nicht immer das gemeldet wird, was tatsächlich vorhanden ist. Es kommt noch dazu - eine Untersuchung, die ich vor vielen Jahren mittlerweile mit einer Kollegin aus dem Weinviertel gemacht habe - dass Allgemeinmediziner regelmäßig keine Diagnosen haben, sondern die haben Verdacht auf eine Gesundheitsstörung, oder die haben ein Bild einer Erkrankung. Nur in 12% oder so waren Diagnosen vorhaben. D.h. das wie Ärzte in ihre eigene Kartei eintragen und was sie auch melden müssen entspricht nicht einer Diagnose.

Ja. Das passt sehr gut zu dem, was wir auch gefunden haben. Ich werde es kurz beschreiben. Wir haben bei 15 Ärzten in verschiedenen Praxen in Großbritannien, ich geb' ihnen kurz etwas, damit Sie sich ein Bild machen können, von der Situation, in der die Daten entstanden sind. Diese 15 Ärzte wurden von zwei simulierten Patientendarstellen besucht, das waren zwei Szenarien.. das eine Szenario, man sieht ihn hier, das Szenario der Halsentzündung, da gings darum, im vergangenen Jahr, mehrmals wiedergekommene, schmerzhafte Halsentzündung, die Stimme ist für den Beruf als Telefonverkäufer wichtig, und er hat die Sorge, dass bei wiederkehrenden Halsentzündungen das provisionsbasierte Einkommen nicht für die Familie reicht. Er hatte als zusätzliches Material den Ausdruck eines geröteten Rachens mit, also als die Ärzte sagten, 'Ok, jetzt würden wir gerne mal reinschauen in den Mund', dann hat er das Bild hergezeigt. Das zweite Szenario, ich glaube hier sieht man sie am Rande, waren Gelenkschmerzen; seit 6 Wochen besonders morgens Gelenkschmerzen und Schwellungen in Händen und Füßen, und sie hat die Sorge, dass sie Arthritis hat, und dass es so ist wie bei ihrer Großmutter, die zunehmend körperlich eingeschränkt worden ist dadurch. Und sie hatte einen Ausdruck einer Abbildung von geschwollenen Finger- und Handgelenken mit. Da war es so, dass wir, ich gebs ihnen einfach so, es ist jetzt recht viel Text, damit Sie einen Eindruck bekommen. Das sind jeweils die Diagnosen, die diese Ärzte produziert haben, wenn sie kurz drüberfliegen.

Ja. ... Ja.

Und man sieht da doch eine gewisse Variabilität in der Diagnose. Und jetzt von diesen verschiedenen Diagnosen her - und Sie sehen auch, wenn nichts dort steht, wurde nichts diagnostiziert, würden Sie sagen, dass dieser

Möglichkeitsraum, dass der erwartbar ist, für den selben Patienten? Und vielleicht noch vorausgeschickt, die simulierten Patienten waren recht gut, in dem Sinn, dass sie vielleicht, mit kleinen Ausnahmen hier und dort, wo sie improvisiert haben, weil es eine sehr spezifische Frage war, wo ihnen nichts eingefallen ist, waren sie relativ einheitlich. Würden Sie sagen, dass das erwartbar ist, dass das die selbe Person ist, dass dieser Möglichkeitsraum an Diagnosen zur selben Person gehören kann?

Ohne weiteres. Das ist eben.. das is ein Grundproblem in der Epidemiologie, dass wir in der Primärversorgung keine Diagnostik zur Verfügung haben, sondern wir haben im Regelfall einen mehr oder weniger gut ausgebildeten und eine mehr oder weniger am Stand des Wissens verbliebene, Ärztliche Person, die versucht, sich einmal ein Bild zu machen. Da es aber für jede Erkrankung - sie sehen hier die Bandbreite von viral bis bakteriell - da diesen Ärzten aber unter diesen Umständen gar nicht das diagnostische Inventar zur Verfügung steht, müssen sie sich auf Bilder von Erkrankungen festlegen, und normalerweise ist es ja ein Herantasten und - vor 30 Jahren oder 40 Jahren hat schon der Braun das war sozusagen der erste.. kennen sie den, der Vater der wissenschaftlichen Allgemeinmedizin, hat damals schon gesagt: "Das Häufige ist häufig", jetzt taste ich mich heran an den geröteten Rachen und denke mir: "in 90% aller Fälle wird das des sein, in 2% aller Fälle wird das des sein, bei zwei weiteren % kann das das sein, und dann hab ich noch ein paar weitere Möglichkeiten, die es wahrscheinlich nicht sind. Also schieß ich jetzt mal auf die 90% mit meiner ersten Therapie , und dann sehen wir weiter. Wenn ich von vornherein annehme, es gibt einen absehbar schweren verlauf der Erkrankung, dann werde ich weitere diagnostische Maßnahmen in Anspruch nehmen, also die Leute überweisen, sie ins Spital einweisen oder dergleichen. Also die Primärversorger sind hier wirklich auf ihr Wissen angewiesen, und auf ihr Wissen angewiesen, und auf ihr Instrumentarium, das sie zur Verfügung haben, in der GP Praxis, und das ist nicht allzu viel. Im Regelfall reichts aber, weil das meiste - das häufige ist häufig - das meiste sind banale Sachen, die leicht beherrschbar sind. aber: zur Auswertung kommen daher im Regelfall keine Diagnosen, sondern eben Bilder von Erkrankungen. sieht aus wie - hat Beschwerden als ob - und das kann ich mit den vorhandenen Kodierungen, den icd10, nicht abdecken. dazu brauch ich andere Kodierungen, den ICPC2, den sie wahrscheinlich kennen, der deckt das, der bildet das sehr gut ab. und ich kann ohne weiteres mit ICPC2 - wir haben damals beim Morbiditätsregister eben den ICPC2 verwendet, mit einigen Erweiterungen, der kann das sehr gut abbilden. ich hab dann halt keine Diagnosen, das macht aber nichts, ich hab trotzdem ein bild der Erkrankungen, der Verbreitung der Krankheiten in der Bevölkerung, das ausreichend ist.

Sehr interessant. Die Unterschied, die wir gesehen haben, waren einerseits, wie sie es beschrieben haben, dass in der Interpretation der Ärzte auch Unterschiede liegen, auch explizit gemacht durch die Ärzte, dass sie eher ganzheitlich das Leiden als Gegenstand der Behandlung sehen, oder eher eine klare Symptomatik, da gibt's auch eben im Fall des Patienten mit Halsentzündung eine ganze Bandbreite von 'Schlafen Sie sich aus', sinngemäß, bis hin zu einer Menge an Untersuchungen. Darüber hinaus war es auch so, dass die

verwendeten Computersysteme auch einen Unterschied produziert haben, in der Hinsicht, dass 1 von den 3 Systemen die Ärzte dazu zwingt, Codes zu verwenden, und ihren Freitext dann entlang dieser Codes zu schreiben, und die anderen Systemen ihnen einerseits ermöglichen, ein paar Codes zu definieren, und dann haben sie ihr Textfeld, wo sie ihren Gedankengang dokumentieren können. Würden Sie sagen, dass in der Epidemiologie bedacht wird, dass verschiedene Systeme auch zu verschiedenen Outcomes führen, unabhängig von dem Kontext der ärztlichen Tätigkeit?

Ich hab an und für sich nichts dagegen, Routine-Daten auszuwerten. Solange sich die Leute, die das auswerten, und die daraus Schlüsse ziehen, der Tatsache bewusst sind, welche Kriterien diese Daten erfüllen. Ich habe viel dagegen, wenn der Hauptverband publiziert, in Österreich gibt es so und so viele Hypertoniker, udn das haben wir errechnet aufgrund der Verschreibungshäufigkeit von Antihypertonika. das ist einfach blödsinn. das ist wirklich Blödsinn. wenn ich aber sage, ich werte diese Daten trotzdem aus, und ich weise darauf hin, dass hier eine gigantische Unsicherheit vorliegt, eben aufgrund der systeme, die zugrunde liegen, die mich zu einem code zwingen, dann ist das durchaus elraubt, wenn darauf hingewiesen wird. der bessere weg ist schon, glaube ich, dass ich entweder ausweiche auf ein Codierungssystem, das solche Unschärfen erlaubt, oder dass ich auf solche Freitextanalysen gehe und dann schau, was haben die wirklich gemeint. und hier ist es halt unter umständen notwendig, dass nachgefragt wird, bei bestimmten Freitexten die nicht verständlich sind, und es wird auch notwendig sein, dass sich das ganze einE MedizinerIn anschaut, weil u.U. das fachliche Wissen notwendig ist, um das beurteilen zu können.

Weil Sie es angesprochen haben: Wie würden Sie sagen, kann man solche Daten sinnvoll validieren? Wenn man jetzt aus einer Epidemiologie-Perspektive sagt, man möchte mit solchen Daten Aussagen treffen, wie sehen Sie die Freitextanalyse z.B., oder Triangulation mit Referral-Briefen an Fachärzte, oder NLP? Was wäre aus ihrer Erfahrung die bessere Methode.

Das ist schwer zu sagen. Es kommt auch auf die Datenmenge an. Bei unserer Studie in Amstetten haben wir das so gemacht, dass wir die Aufnahmediagnosen - ich verwende das Wort ungern, weil es eben keine Diagnosen waren, sondern es waren Beschwerdebilder, die die Patienten geschildert haben,d as haben wir von 2 unabhängigen, niedergelassenen Ärzten begutachten lasen, und wir haben nachher geschaut, wie gut die übereinstimmen, und die haben zu gut 2/3 übereingestimmt. das wäre eine Möglichkeit, die mir einfällt. ich vertraue hier eher nicht der Maschine, würde ich sagen, weil gerade in der Medizin so viele Synonyme vorhanden sind, und so viele Nebenbefunde, die aber in eine abschließende Beurteilung nicht einfließen - vorhanden sind, dass ich mir schwertue, eine Maschine das machen zu lassen.

Im Freitext.. wir haben da verschiedene Verwendungsformen unterscheiden können. Das war z.B. um das Narrativ des Patienten zu dokumentieren, es war auch ein Versuch, eine Strategie, um sich gegen irgendwelche negativen Folgen abzusichern, andererseits gings auch darum, den Patienten -

wenn Ärzte das System quasi gemeinsam verwendet haben, wenn es in einem Gespräch quasi darum ging, einmal zu etablieren, worum es geht, und dann machen sie es explizit, indem sie es gemeinsam ins System hineinschreiben, das war auch eine Verwendungsform. Das heißt es gibt unterschiedliche Perspektiven in den Daten, davon abhängig, wo sie herkommen, ob Sie von Arzt und Patient gemeinsam oder nur dem Arzt herkommen. Außerdem auch um gewisse Anker der Erinnerung zu platzieren, z.B. ein Kind, das besonders freundlich gelacht hat, hat auch ein Arzt als Beispiel gegeben. Um so eine spezielle Note da drin zu haben, wenn sie selber wieder lesen, damit sie sich erinnern können. Wäre es aus epidemiologischer Sicht interessant, verschiedene dieser Faktoren besser sichtbar zu machen, z.B. die Art und Weise, wie Patienten Dinge sehen, vielleicht auch gegenüberzustellen, wie Ärzte ihr Resultat sehen. Wäre das für einen Epidemiologen interessant?

Das ist aus meiner Sicht sehr interessant. Allerdings es kommt auf das Ziel der Studie an. Ich hab in der Medizin ja ein unheimlich breites feld, das bearbeitet wird, und es geht einerseits um sehr ernste Erkrankungen wo es darum geht, möglichst schnell zu einer Diagnose zu kommen und eine Therapie zu beginnen und andererseits, am anderen Ende der Skala, hab ich eine psychologische Betreuung von nicht krankheitswertigen Beschwerden, ich hab jetzt grad einen arikel gelesen über das PMS, es gibt ja ein Beschwerdebild des noch stärker , schwerer ist und Behandlung bedarf, während das PMS von den Allermeisten abgetan wird als 'naja, wenns nicht allzuschwer ist, dann is es am besten, wenn die Frauen es ertragen' es sind halt irgendwas von 2 Tagen bis 2 Wochen pro Monat wo die Frauen sich schwer tun, und wenn alle eingebunden sind und Bescheid wissen, dann ist das an und für sich nicht krankheitswertig zu verstehen und dann kann ich die Frau beraten im Hinblick darauf, dass sie gewisse Nahrungsmittel vermeidet, oder zusätzlich zuführt, oder irgendwie a bissl so.. in die richtung leichte, psychotherapeutische Intervention geht. und da ich diese enorme Breite der Beschwerden habe, kommt es sehr darauf an, was dem zugrunde liegt, wonach sie jetzt gefragt haben. es kann irrsinnig interessant sein, hier die sicht die Patienten der sicht der behandelnden gegenüberzustellen, und auf der anderen seite kanns auch völliger Unsinn sein, weil das eben eine hochspezialisierte, oder ganz genau definierte Erkrankung ist, wo die Meinung des Patienten uninteressant is, weil wir wissen was das is und das gehört so und so behandelt und da gibts überhaupt keine Diskussion.

Alles klar.

Wichtig ist aber schon, in der arbeit der GPs.. aus psychohyigenischen gründen, die sitzen ja jeden tag in ihrer ordination, hier sind solche Aufzeichnungen wahrscheinlich schon sehr wichtig für die Psychohygiene, dass die überhaupt die Tage überstehen, dass sie so, highlights sich aufschreiben, wenn sie wieder drübergehen, sich erinnern, das is glaub ich schon sehr wichtig. wir dürfen nicht glauben, dass die Aufzeichnungen der Ärzte nur dazu dienen, Diagnosen auszuwerten oder herauszusaugen, e sgeht da um viel viel mehr. und wenn wir nur die Diagnosen herausnehmen tun wir denen eigentlich unrecht, weil da is vü mehr drinnen, eben solche Atmosphärischen sachen. ich denk gerade an

Kinderärzte z.B., wo es irrsinnig wichtig, ob die Kinder ununterbrochen heulen oder eben lachen.. und man hat gleichzeitig ein Gefühl dem Kind gehts in der Familie gut und da brauch ich mir keine Sorgen machen.

Diese Studie, die sie erwähnt haben, würde mich interessieren. Gibt es dafür vielleicht Literatur?

Ich schick ihnen das einfach!

Super danke. Das wars soweit von meinem Plan. Gibt es vielleicht noch Dinge, die Sie wichtig fanden, die ihnen noch einfallen würden?

Also ich weiß nicht, wie sehr Ihre Arbeit in Richtung Auswertung von ELGA-Daten geht.

In erster Linie hat es nichts zu tun. Es ist aber ein interessanter Vergleich.

Ich halte es für sehr gefährlich, diese Daten auszuwerten. Ich hab da einen Vortrag über die Qualität epidemiologischer daten im Hauptverband gehalten, im Prinzip haben wir das selbe geredet wie wir jetzt, ja, worums halt bei der Qualität von epidemiologischen Daten geht, ich halte es nämlich für sehr gefährlich, die ELGA-Daten für sowas heranzuziehen. auch hier kommts wieder drauf an, wenn ich von vornherein a Diagnose hab, is alles klar, aber wenn ich Beschwerdebilder auswerte, dann kommts ja auch unheimlich drauf an, wie nennen das Ärzte in ihrer eigenen Ordination, wir haben uns.. der arme Kollege, der is immer noch nicht fertig mit seiner Diplomarbeit. Der ist Monate gesessen um Synonyme auf einen Nenner zu bringen. weil eben die Leute, die in Ambulanzen gekommen sind, das is einfach so hingeschrieben wroden die Beschwerdebilder.. alle diese Synonyme sind sselbstverständlich in ELGA drinnen. ich halte es für falsch, solche Dinge ungefiltert auszuwerten. Und damals vor 2 Jahren, hab ich ihnen das eben gesagt, und da is auch einer vom Hauptverband, der eben hauptverantwortlich is für ELGA.. is auch aufgestanden und hat gesagt, es is sicher nicht gedacht, die ELGA-Daten auszuwerten!

Ok.

Und ich hab ihm damals schon gesagt.. wenn Sie das so öffentlich sagen, es freut mich sehr, ich hab damals schon gezweifelt von dieser Aussage.. ich merk wie von verschiedenen Seiten die Forderung kommt, die ELGA-Daten auszuwerten. Ich glaube wir produzieren hier wirklich einen haufen Mist.

Weil die Daten zu wenig auswertbar sind, weil sie zu wenig verstanden werden?

Ja. Weil jedes zweite nicht eingetragen wird, weil Leute den Aufwand scheuen, das System nochmal.. dann fällt ihm im nachhinein noch was ein, der Patient is aber schon weg. Natürlich ergänzt er das in seiner kartei, aber nicht in der ELGA. der ärztliche Beruf besteht ja nicht nur im Patientenkontakt, sondern auch im Nachdenken über den Fall, hab ich da irgendwas vergessen. und das alles is nicht drinnen in der ELGA.

Wäre ein System, das weniger Einzelbilder speichert, sondern mehr den Prozess abbilden kann, sondern auch den Prozess im Allgemeinen, wäre das besser?

Ja, das halte ich für möglich, mit der gebotenen Vorsicht natürlich. und genau in die richtung ist unser Morbiditätsregister damals gegangen.

Ah, interessant.

Dass wir uns überhaupt nicht auf Diagnosen stürzen, sondern dass wir genau in diese richtung gehen. wWr wollten ein bild der Bevölkerung haben, das sich aus einzelnen Arztbesuchen zusammensetzt. und dann sieht man, wie halt der 30-jährige beginnt, übergewichtig zu werden, und dann mit 38 das erste Mal eine Hypertonie-Diagnose hat, und es so weitergeht. Das wollten wir eigentlich. Wär auch gegangen mit dem System, wir haben aber kein Geld mehr gehabt.

Verstehe. Und dieses System war quasi ein Auswertungsansatz für Daten?

Das Ziel war eine Erfassung der Morbidität zu schaffen, die unabhängig ist von Störfaktoren, nämlich z.b. der Störfaktor der Krankenversicherung oder der Deckelung der Ärzte bei verschiedenen Diagnosen und all dem, sondern das Ziel war, wie ist das Bild der Allgemeinedziiner in ihrer Praxis, was erleben sie so tag-täglich, abgebildet in Form von Patienteakten. wir wollten halt auf 90 Ärzte aufstocken in Österreich, dann hätten wir eben, dann hätten wir die ganzen Epidemien abgebildet gehabt, dann hätten wir Übergewicht, Risikofaktoren, Rauchen, Alkoholkonsum.. dann hätten wir das abgebildet gehabt in einer weise, dem ich sehr vertraut hätte.

Weil das diesen Blick in die Länge hat und nicht so diese Einzelne?

Genau und weil es nicht beeinflusst ist durch Abrechnugsdaten, durch Remuneration, wieviel krieg ich wofür durch die Versicherung.

Das haben wir auch beobachtet, dass Ärzte Dinge dokumentiert haben, wo ihnen in ihrem Praxisverbund von einem verantwortlichen Manager gesagt worden ist, das ist wichtig, dass das im System drin steht, weil das für einen Audit wichtig ist.

Ich unterstelle den Ärzten überhaupt keine böse Absicht, ich weiß einfach, dass die Abrechung die Dokumentation beeinflusst, da gibt es überhaupt keinen Zweifel. Ich kann sie trotzdem auswerten, ich muss aber darauf hinweisen, dass dieser Verzerrungsfaktor enthalten ist. Das ist es eben.

Und diese Verzerrung, können Sie beschrieben, welche Effekte diese hat?

Es geht um das Einkommen der Ärze. Im anderen bereich, im Spitalsbereich, wir wissen ja ganz eindeutig, dass das LKF-System so gebogen wird, dass möglichst viele Punkte für die Spitäler erzielt werden. und da gibts Computerprogramme, die für die einzelnen p Patienten vorschlagen, die Untersuchung noch zu machen, und in die Richtung zu gehen, weil es dann mehr Punkte zum Lukrieren gibt.. sobald die wirtschaftliche Situation hineinspielt, werden Sachen dokumentiert, die nicht ganz der Wahrheit entsprechen.

Ok, danke. Das wärs von meiner Seite. Haben Sie noch Kommentare? Sonst möchte ich vielen Dank sagen für Ihre Zeit und Ihren wertvollen Beitrag. Ist sehr spannend für uns.

Ja, irrsinnig interessant.. so interessant. ich kenn einige wenige Ärzte, die beschlossen haben, nicht viel Geld verdienen zu wollen. und die behandeln die Leute, vollkommen anders in der Ordination. Die nehmen sich Zeit, die setzen sich hin. die Leute wissen "die will wirklich wissen, was ich hab", das wird manche auch abhalten hinzugehen. Ich muss nochmal den Braun ins Spiel bringen.. der hat zwei Probleme gehabt.. 1) Zu wenig Zeit für seine Patienten, weil bei ihm das Wartezimmer voll is.. deswegen musste er seinen Arbeitsablauf optimieren.. und 2) weil er sich hignelegt am Abend , sind ihm immer wieder die Patienten durch den Kopf gegangen. und er hat sich immer wieder die Frage stellen müssen, um Gottes willen, hoffentlich hab ich nicht etwas übersehen. also dieser abwendbar gefährliche Verlauf einer Erkrankung. und die zwei Sachen wollte er angehen..und er hat so Formulare entworfen.. für verschiedene Krankheitsentitäten.. und anhand dieser Formulare hat er erstens nichts vergessen, udn zweitens war er viel schneller im Ablauf. Der war genial, der Brown. z.b. diese eine Kollegin, die Dr. Fink, die nimmt sich Zeit, der is es wurscht, wieviel Geld sie verdient. Und so macht man die bessere Medizin, aber in den meisten Fällen ist man ökonomischen Zwängen ausgesetzt.

Sehr interessant. Das wars soweit von meiner Seite. Haben Sie sonst noch Gedanken, ist noch etwas offen? Sonst möchte ich sagen: Vielen Dank für Ihre Zeit!

Gerne.

B.3 Consent Form

Study Description (Excerpt)

Ich beschäftige mich in meiner Masterarbeit (Medieninformatik an der Technischen Universität Wien) mit Aspekten der Sekundäranalyse des elektronischen Gesundheitsaktes.

Konkret untersuche ich Unterschiede in der Verwendungspraxis von Codes und freiem Text durch praktische Ärzte und welche Implikationen sich darin für Sekundäranalysen und in weiterer Folge auch aus Sicht z.B. der Epidemiologie für die Gestaltung von Informationstechnologie für ÄrztInnen erkennen lassen. Dafür habe ich im Rahmen einer qualitativen Videoanalyse (zwei simulierte PatientInnen mit Szenario »Gliederschmerzen« und »Halsentzündung« in Praxen von 15 praktischen Ärzten) vielfältige Unterschiede in der resultierenden Diagnose, Behandlung und Verwendung des elektronischen Gesundheitsaktes dokumentieren können. Um die Bedeutung dieser Unterschiede in der Dokumentation für Sekundäranalysen besser verstehen zu können, würde ich sehr gerne ExpertInnen der Epidemiologie befragen.

Besonders interessant wären dabei ihre Erfahrungen mit Sekundäranalysen, Ihr Forschungsinteresse bzgl. der die ärztliche Behandlung dokumentierenden Daten sowie etwa auch, was aus Ihrer Sicht besonders wichtige Qualitätsmerkmale dieser Daten sind. Ihr Input in diesen Aspekten ist enorm wertvoll für die weitere Verbesserung von Computersystemen für Ärzte und die Nutzbarmachung der Daten für medizinische Forschung.

Consent Form

Experteninterview – Einverständniserklärung

- Inhalt des Interviews ist Ihre Erfahrung mit epidemiologischen Studien und den dafür verwendeten Daten. Wir sind dazu nicht "kritisch", bzw. es gibt keine richtigen/falschen Antworten. Sie können die Beantwortung einzelner Fragen ohne Angabe von Gründen ablehnen und Ihre Teilnahme jederzeit beenden.
- Das Interview wird zur anonymisierten Transkription aufgezeichnet. Die Audioaufzeichnung wird niemandem sonst zugänglich gemacht und nach erfolgter Transkription und Analyse gelöscht.

Dieses Interview stellt einen Teil der praktischen Arbeit im Rahmen meiner Diplomarbeit dar. Für Rückfragen, Beschwerden und sonstige Kommentare steht meine Betreuerin, Prof. Geraldine Fitzpatrick, PhD unter <E-Mail-Adresse> erreichbar.

Vielen herzlichen Dank für Ihre Teilname!
Christian Löw

Ich bin damit einverstanden, dass das Interview aufgezeichnet wird.

[ ] Ja / [ ] Nein

Ich bin damit einverstanden, dass wörtliche, anonymisierte Zitate in der Diplomarbeit oder weiterführenden Publikationen verwendet werden.

[ ] Ja / [ ] Nein

<Unterschrift>

APPENDIX C

Appendix: Compilation of EHR Usage Charts

This appendix section provides all charts used in chapters 5 and 6 in a compiled section to make comparison easier.

C.1 Category Usage Schematics

Figures C.1, C.2 and C.3 contain the category usage schematic for EMIS LV, SystmOne and Vision.

C.2 Word Counts

Figures C.4, C.5 and C.6 depict word count with EMIS LV, SystmOne and Vision. Figure C.7 yields an overview of EHR documentation word count across surgeries.

C.3 Code and Category Usage

Figures C.8 and C.9 yield an overview of codes and categorization used by GPs across surgeries.

C.4 Word Count and Age

Figure C.10 yields a comparison of word count in the EHR documentation and age range of the GP. Figure C.11 provides an overview of GPs and surgeries.

Figure C.1: Usage of EHR categories in EMIS LV with *Scott's* case (bottom) and *Jane's* case (top) with respective code count (outter right column). <>.. Free-Text, **.. Code, ||.. mentioned in consultation.

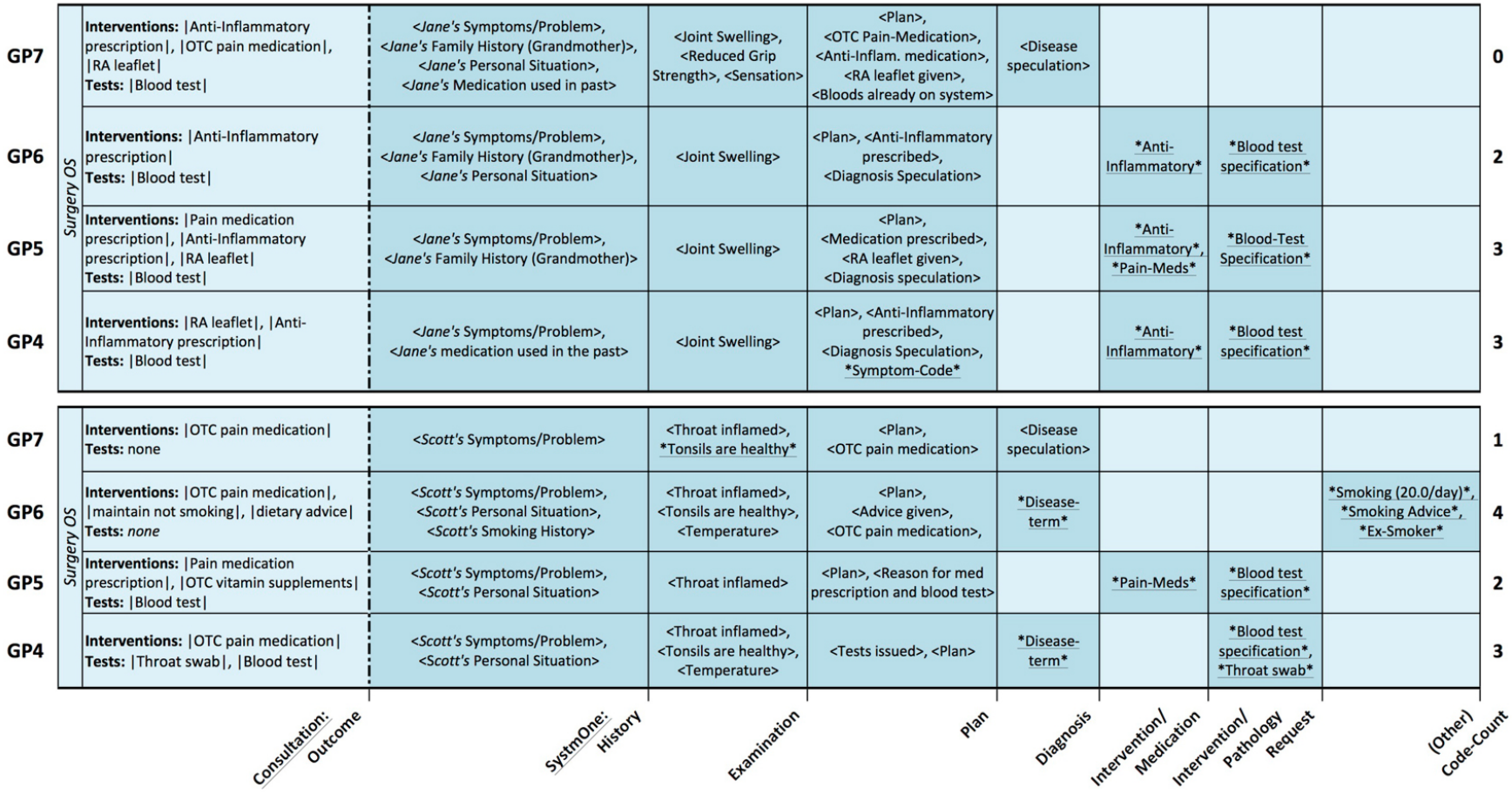


Figure C.2: Usage of EHR categories in SystmOne with *Scott's* case (bottom) and *Jane's* case (top) with respective code count (outter right column). <>.. Free-Text, **.. Code, ||.. mentioned in consultation.

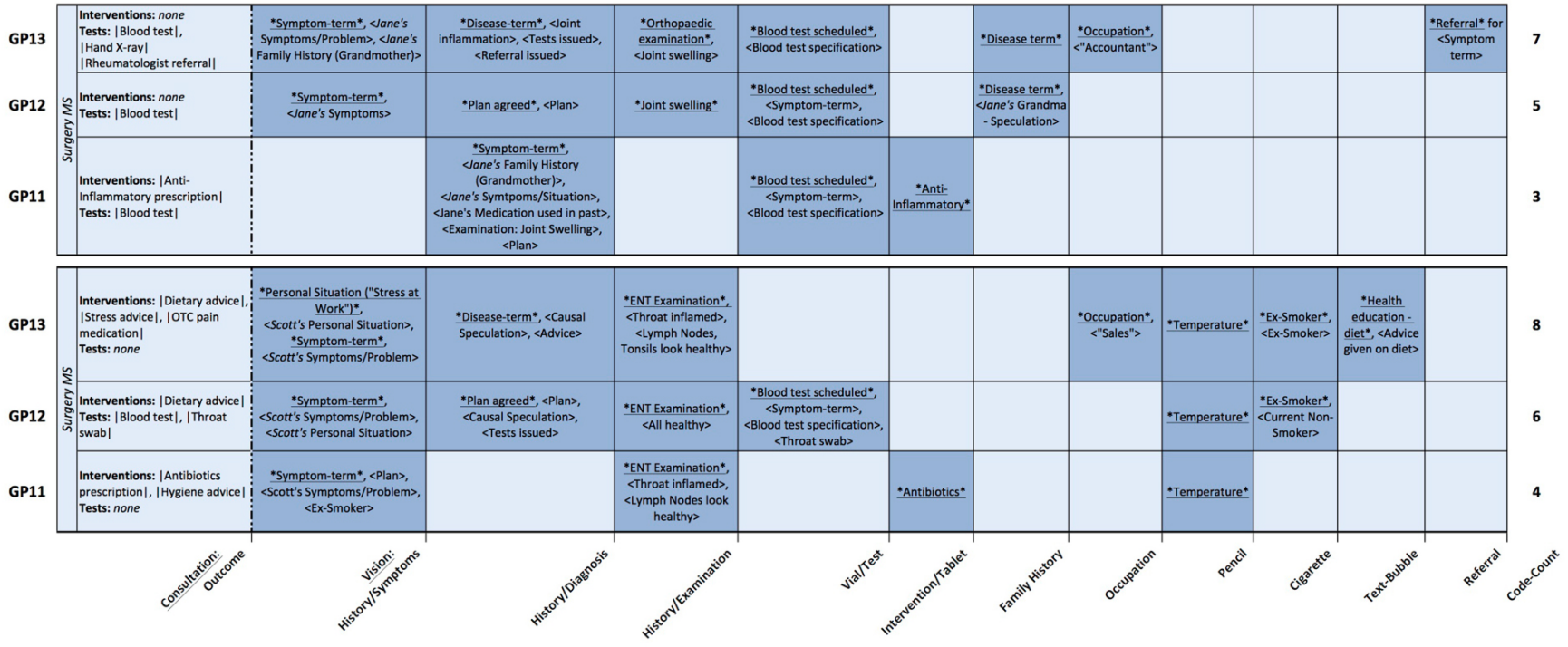
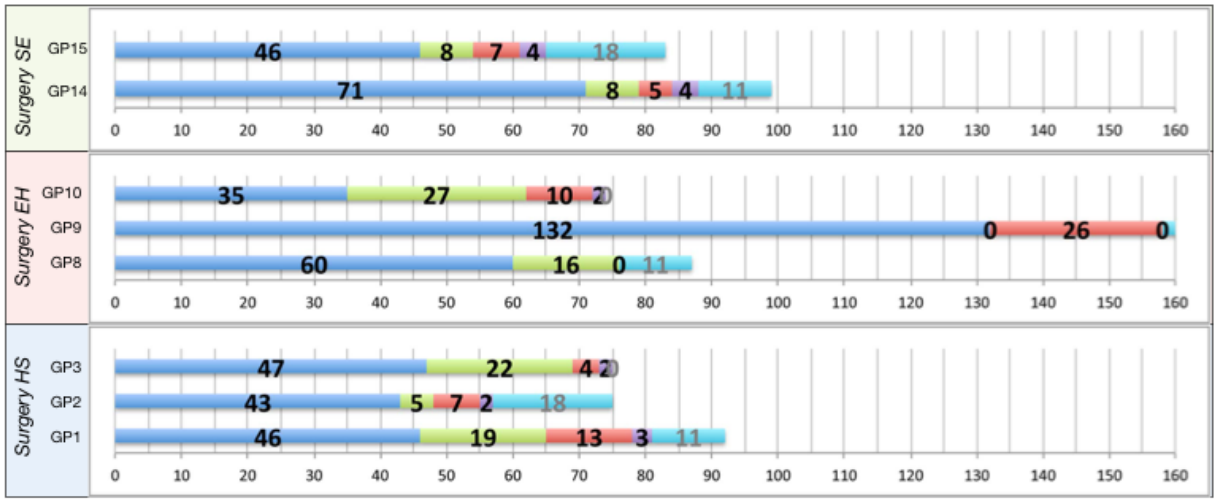
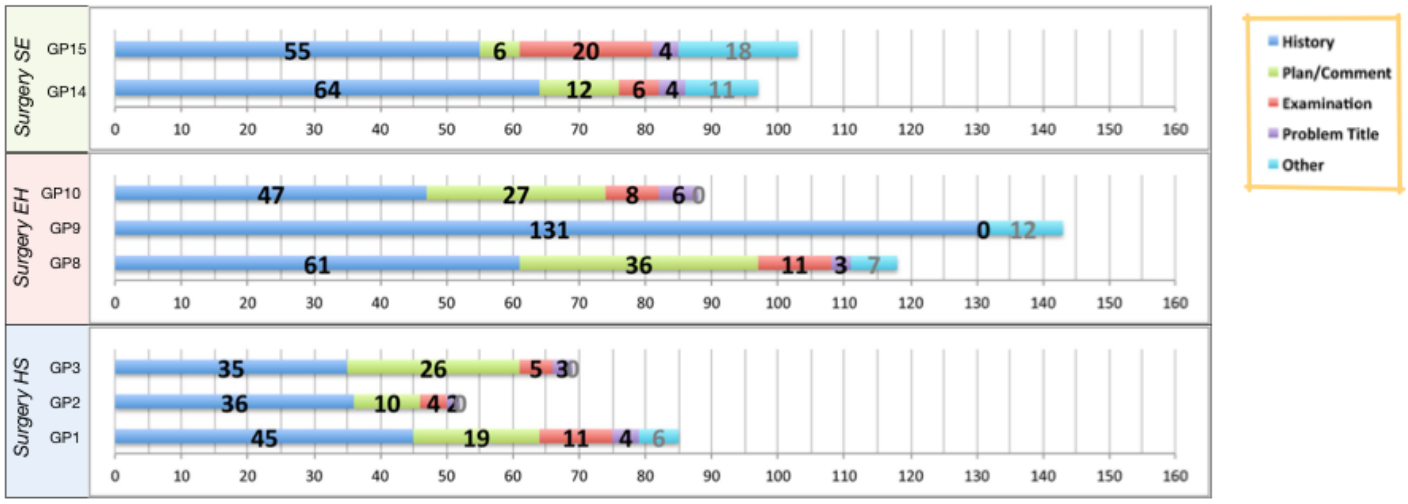


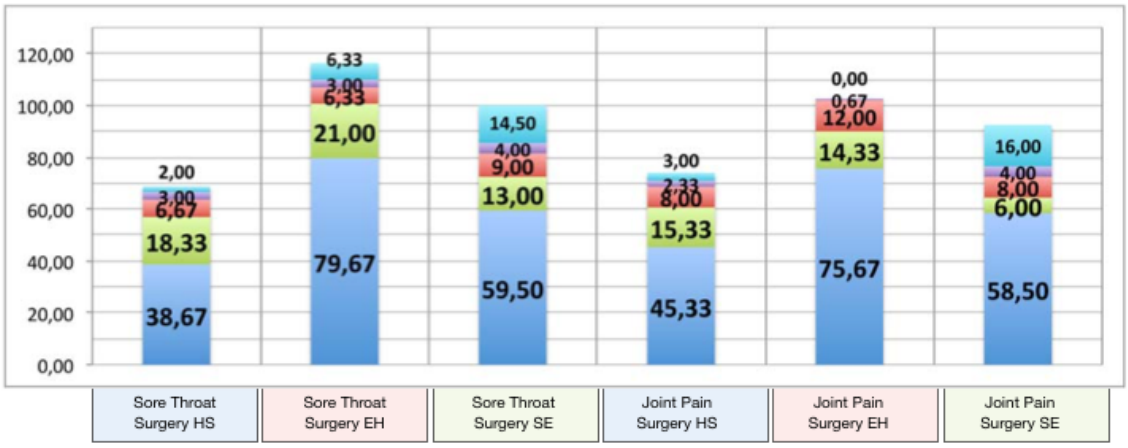
Figure C.3: Usage of EHR categories in Vision with *Scott's* case (bottom) and *Jane's* case (top) with respective code count (outter right column). <>.. Free-Text, **.. Code, ||.. mentioned in consultation.



(a) Joint Pain (JP): Category word count in EMIS LV



(b) Sore Throat (ST): Category word count in EMIS LV



(c) Average category word count in EMIS LV

Figure C.4: Word count in EHR documentation within EMIS LV.

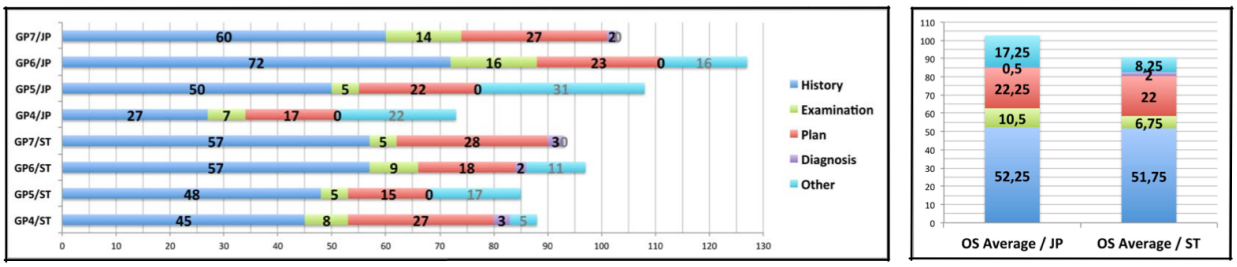


Figure C.5: Word count of free text categories in SystmOne/surgery OS.

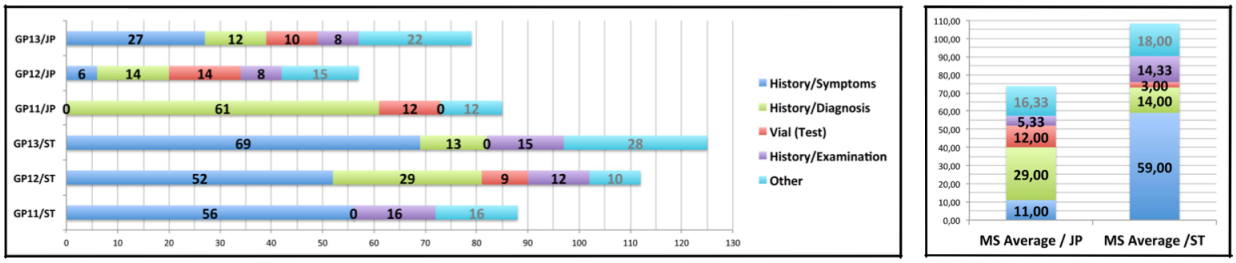


Figure C.6: Word count of free text categories in Vision/surgery MS.

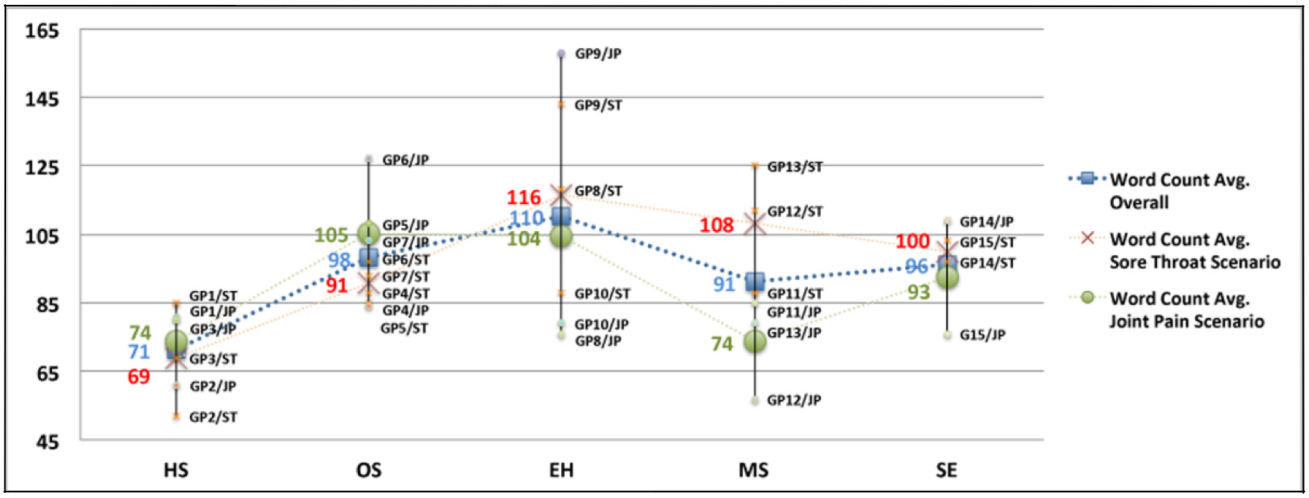


Figure C.7: Overview of EHR documentation word count across surgeries. EMIS LV (Surgeries HS, EH and SE), SystmOne (Surgery OS) and Vision (Surgery MS).

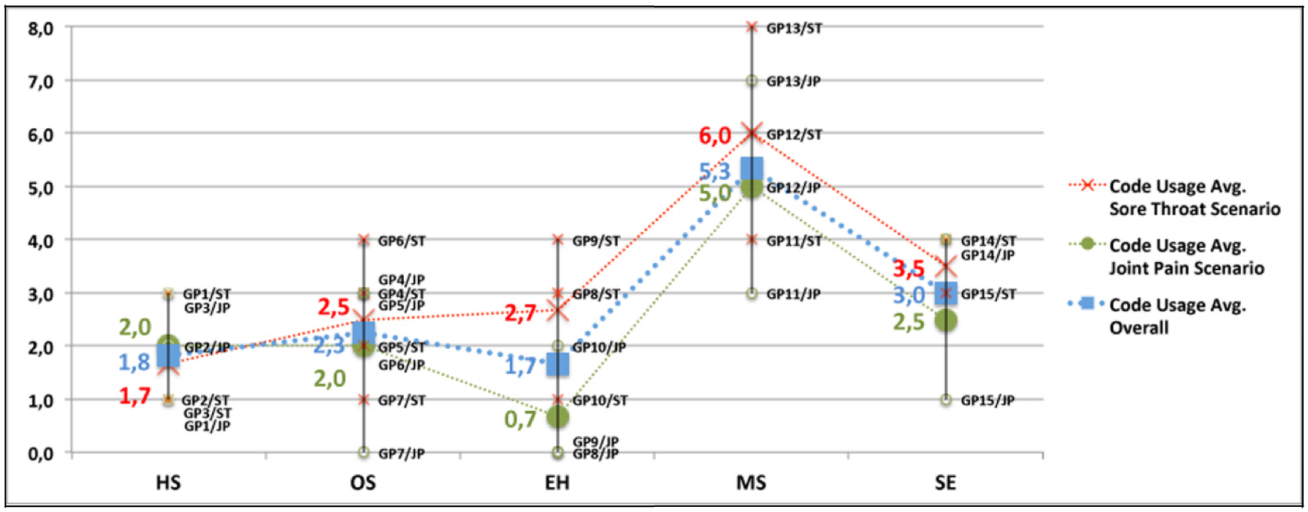


Figure C.8: Overview of amount of codes used in EHR documentation across surgeries. EMIS LV (Surgeries HS, EH and SE), SystmOne (Surgery OS) and Vision (Surgery MS).

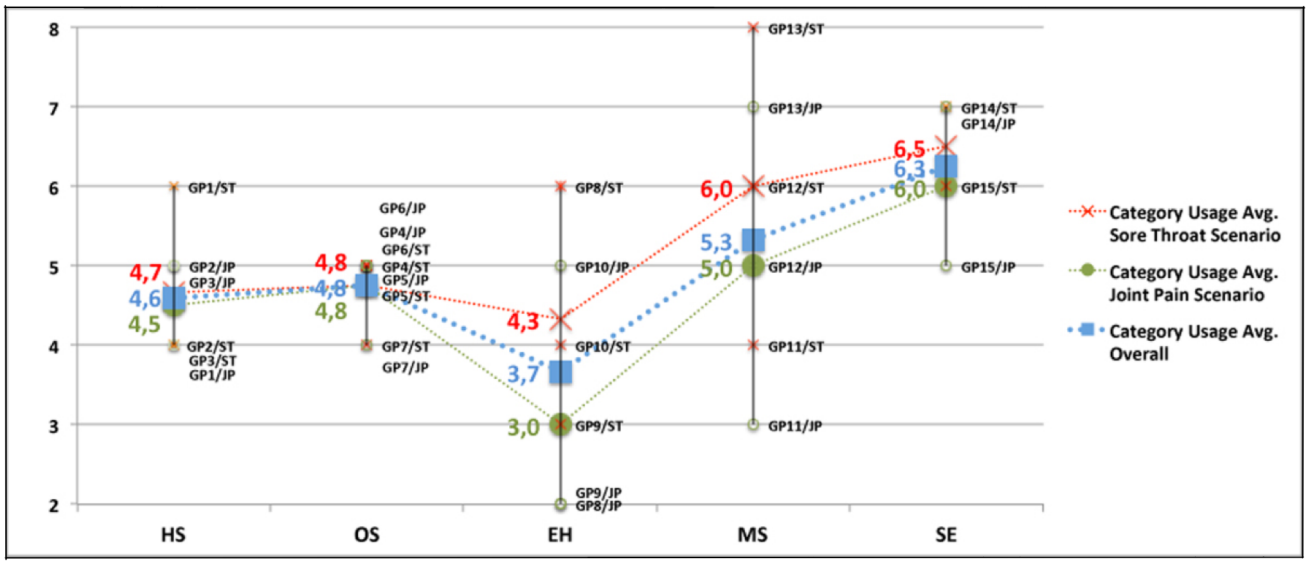


Figure C.9: Overview of categorization of EHR documentation across surgeries. EMIS LV (Surgeries HS, EH and SE), SystmOne (Surgery OS) and Vision (Surgery MS).

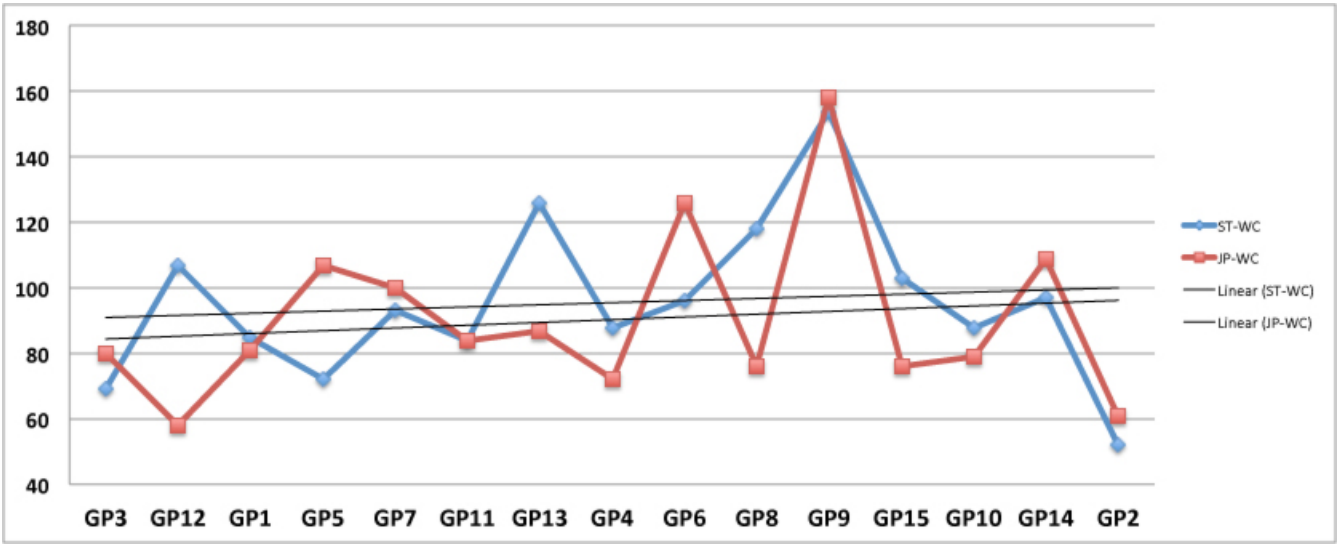


Figure C.10: Word count with Jane (JP-Scenario) and Scott (ST-Scenario). GPs in order of age. An overview of the GPs with gender and age range can also be found in figure C.11.

Surgery	EHR System	GPs (gender, age range)			
Surgery HS	*EMIS LV*	**GP1** female, 30s	**GP2** male, 60s	**GP3** female, 20s	
Surgery OS	*SystmOne*	**GP4** male, 40s	**GP5** female, 30s	**GP6** female, 40s	**GP7** female, 30s
Surgery EH	*EMIS LV*	**GP8** female, 40s	**GP9** male, 40s	**GP10** male, 50s	
Surgery MS	*Vision*	**GP11** female, 30s	**GP12** female, 30s	**GP13** female, 20s	
Surgery SE	*EMIS LV*	**GP14** female, 50s	**GP15** female, 40s		

Figure C.11: Overview of GPs and surgeries: GPs with gender and age range, surgeries with employed EHR system.

List of Figures

1.1 Usage of EHRs among primary actors (e.g. GP, colleagues, specialists, ..) is embedded in collaboration that, at least to some part, can rely on the „ground truth“ of patient care to negotiate and validate its process. Secondary Analysis is disjunct from this system and often has to make assumptions on how the collaborative artefacts that constitute EHR documentation are to be interpreted. 5
1.2 Outline of the several tasks and related results. 7

2.1 Study situation: Consultations were preceded and followed by prep time the GPs used to look at the simulated patient's records and document the consultation. 13
2.2 ST scenario persona named Scott Marshall describes his symptoms (left), JP scenario persona named Jane Hallam has 'her' hand examined (right). 15

3.1 Perspectives can be deceiving. (Source unknown.) 20
3.2 Overview of GPs and surgeries: GPs with gender and age range, surgeries with employed EHR system. 23
3.3 One recording session with a GP consists of the Joint Pain and Sore Throat (**JP**, **ST**) scenario consultations and a walkthrough interview afterwards. . . 24
3.4 Overview of the analysis process applied to the data set. Three passes of viewing the consultations and interviews yielded coding schemes in different categories towards themes in line with a grounded theory approach. 27
3.5 Progression of the analysis process towards theme refinement. 28
3.6 List of experts interviewed in the course of this work and their reference (E1-E6) in the thesis. 30

4.1 *Jane*'s medical history placed in the EHR systems prior to the consultations. (GP11. Surgery MS. EHR-System: Vision.) 35
4.2 GPs attention on the actor's hands as she describes her symptoms (left), Examination with print-out: "What happens when I squeeze there?" (right). GP1. 39
4.3 Basic Focal Themes in Consultations. 40
4.4 Overview of consultation outcomes with *Scott* / The Sore Throat scenario. . 46
4.5 Overview of consultation outcomes with *Jane* / The Joint Pain scenario. . . . 47

5.1 The interface of EMIS LV. (a) Overview, (b) GP timetable, different patient overviews (c), (d). 53
5.2 GP14 documents *Scott*'s case (who, in this case, was created as "Mr. Donald Duck". 54
5.3 GP14 documents *Scott*'s case (who, in this case, is named "Mr. Donald Duck". 55
5.4 Patient journal overview with SystmOne. 56
5.5 SystmOne consultation data entry interface form (top), being filled out (middle) and how it is listed in the patient's consultation journal (bottom). 57
5.6 Predictive code suggestion in SystmOne. 58
5.7 Patient Overview with Vision. Vision offers a priority setting in addition to the problem label as structuring elements for summaries (arrows). 59
5.8 Documenting with Vision starts with entering a Read Code. 60
5.9 Entering Family History with Vision. 61
5.10 A symptom-related Read Code yields the 'History - Add' interface. Vision. . 61
5.11 A procedure-related Read Code yields a specific, test-related interface. Vision. 62
5.12 A finished EHR consultation entry of *Scott's* case (GP12). Yellow highlighting signifies codes. Vision. 63
5.13 Usage of EHR categories in EMIS LV with *Scott's* case (bottom) and *Jane's* case (top) with respective code count (outter right column). <>.. Free-Text, **.. Code, ||.. mentioned in consultation. 66
5.14 Word count in EHR documentation within EMIS LV. 69
5.15 Content landscape of SystmOne EHR documentation of *Scott*'s (bottom) and *Jane*'s case (top). <>.. Free-Text, **.. Code, ||.. mentioned in consultation. 70
5.16 Word count of free text categories in SystmOne/surgery OS. 72
5.17 Content landscape of SystmOne EHR documentation of *Scott*'s (bottom) and *Jane*'s case (top). <>.. Free-Text, **.. Code, ||.. mentioned in consultation. 73
5.18 Word count of free text categories in Vision/surgery MS. 75
5.19 Overview of amount of codes used in EHR documentation across surgeries. EMIS LV (Surgeries HS, EH and SE), SystmOne (Surgery OS) and Vision (Surgery MS). 75
5.20 Overview of categorization of EHR documentation across surgeries. EMIS LV (Surgeries HS, EH and SE), SystmOne (Surgery OS) and Vision (Surgery MS). 76
5.21 Overview of EHR documentation word count across surgeries. EMIS LV (Surgeries HS, EH and SE), SystmOne (Surgery OS) and Vision (Surgery MS). 77

6.1 QOF flags. SystmOne supports highlightig QOF-relevant codes. 83
6.2 The computer as a handy companion in GP practice. 89
6.3 Word count with Jane (JP-Scenario) and Scott (ST-Scenario). GPs in order of age. An overview of the GPs with gender and age range can also be found in figure 3.2. 92

7.1 Usage of EHRs among primary actors (e.g. GP, colleagues, specialists, ..) is embedded in collaboration that, at least to some part, can rely on the „ground truth“ of patient care to negotiate and validate its process. Secondary Analysis is disjunct from this system and often has to make assumptions on how the collaborative artefacts that constitute EHR documentation are to be interpreted. 101
7.2 Different notions of data quality in Electronic Health Records, as found by Häyrinen et al. [47], Weiskopf and Weng [10], Botsis et al. [12], Herrett et al. [46] and Khan et al. [9]. The latter two focused on data quality for coded diagnoses, while the others on EHR data in more general respect. *).. Häyrinen et al. found further quality measures which other authors would, explicitly and implicitly, subsume mostly under *accuracy*. 103
7.3 Common validation strategies with EHR data. CPRD-Standards intrinsic to the database (top), Strategies of validating EHR content in general [10] (middle), Strategies to validate diagnoses [46] (bottom). Dotted lines represent conceptual relationship. 105
7.4 Flow of data from its origin in general practice consultations over data aggregation efforts to secondary analysts. 108

8.1 Surgery MS (Vision) exhibits higher code usage than other surgeries with other EHR systems. EMIS LV (Surgeries HS, EH and SE), SystmOne (Surgery OS) and Vision (Surgery MS). 123
8.2 Overview of categorization of EHR documentation across surgeries. 124
8.3 Comparing the interface of documentation structure with Vision (top) and SystmOne (bottom). 125

C.1 Usage of EHR categories in EMIS LV with *Scott's* case (bottom) and *Jane's* case (top) with respective code count (outter right column). <>.. Free-Text, **.. Code, ||.. mentioned in consultation. 152
C.2 Usage of EHR categories in SystmOne with *Scott's* case (bottom) and *Jane's* case (top) with respective code count (outter right column). <>.. Free-Text, **.. Code, ||.. mentioned in consultation. 153
C.3 Usage of EHR categories in Vision with *Scott's* case (bottom) and *Jane's* case (top) with respective code count (outter right column). <>.. Free-Text, **.. Code, ||.. mentioned in consultation. 154
C.4 Word count in EHR documentation within EMIS LV. 155
C.5 Word count of free text categories in SystmOne/surgery OS. 156
C.6 Word count of free text categories in Vision/surgery MS. 156
C.7 Overview of EHR documentation word count across surgeries. EMIS LV (Surgeries HS, EH and SE), SystmOne (Surgery OS) and Vision (Surgery MS). 156
C.8 Overview of amount of codes used in EHR documentation across surgeries. EMIS LV (Surgeries HS, EH and SE), SystmOne (Surgery OS) and Vision (Surgery MS). 157

C.9 Overview of categorization of EHR documentation across surgeries. EMIS LV (Surgeries HS, EH and SE), SystmOne (Surgery OS) and Vision (Surgery MS). 157
C.10 Word count with Jane (JP-Scenario) and Scott (ST-Scenario). GPs in order of age. An overview of the GPs with gender and age range can also be found in figure C.11. 158
C.11 Overview of GPs and surgeries: GPs with gender and age range, surgeries with employed EHR system. 158

List of Tables

2.1 "Persona variables" of the two scenarios. [17, Table 1] 14

4.1 Persona variables of the two scenarios. [17, Table 1] 33

A.1 Table 138