Towards formal medical reporting

An evaluation in endoscopy

This research was financially supported by Glaxo B.V.
Glaxo B.V. is also gratefully acknowledged for their
financial support for the publication of this thesis.

Towards formal medical reporting

An evaluation in endoscopy

(Naar formele medische verslaglegging
Een evaluatie in de endoscopie)

PROEFSCHRIFT

Ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam op gezag van
de Rector Magnificus
Prof. Dr. P.W.C. Akkermans M.A.
en volgens besluit van het college voor promoties.

De openbare verdediging zal plaats vinden op
woensdag 4 oktober 1995 om 15.45 uur

door

Peter Willem Moorman

geboren te Diepenveen

Promotiecommisie

Promotores:	Prof. Dr. Ir. J.H. van Bemmel Prof. J.H.P. Wilson
Co-promotor:	Dr. A.M. van Ginneken
Overige leden:	Prof. Dr. A.J. Man in't Veld Prof. Dr. P.F. de Vries Robbé Prof. Dr. G.N.J. Tijtgat

Contents

Dankwoord

Uiteraard moet men bij het schrijven van een proefschrift een zo groot mogelijke zorgvuldigheid betrachten. In het geval van een dankwoord impliceert dit volledigheid. Aangezien dat tot een dankhoofdstuk zou hebben geleid beperk ik mij hier tot slechts enkele personen. Diegenen die hier niet genoemd worden zijn zeker niet vergeten.

Allereerst wil ik Jan van Bemmel, Paul Wilson en Mark van Blankenstein bedanken voor de mogelijkheid die zij mij hebben gegeven dit onderzoek te doen. Altijd vond ik hen bereid mee te denken en constructief commentaar te leveren.

Voor veel studies was ik afhankelijk van de medewerking van endoscopisten. Daarom wil ik P. Batenburg, A. Bauer, R. Beukers, H. van Buuren, J. Dees, H. van Essen, H. Geldof, W. Lesterhuis, R. Ouwendijk, J. Scherpenisse, A. van Tilburg, P. Fockens en K. Huibregtse bedanken voor de kostbare tijd die zij in dit onderzoek hebben willen stoppen.

Dit laatste geldt ook voor Johan van der Lei, die altijd bereid was mee te denken en mijn teksten 'af te schieten'.

Mijn grootste dank gaat uit naar Astrid van Ginneken en Peter Siersema. Zonder jullie begeleiding was dit proefschrift er niet gekomen. Hopelijk kijken jullie met net zoveel voldoening terug op de afgelopen periode als ik.

Science is not 'organized common sense';
at its most exciting,
it reformulates our view of the world
by imposing powerful theories against
the ancient prejudices that we call intuition.

Stephen Jay Gould
in
Ever since Darwin

Chapter 1

General introduction

Introduction

For clinical practice, the patient record is the principal repository for information concerning a patient's health care [1]. For centuries, medical notes were brief comments used by their author to trigger a fuller recollection of his patients. In the late nineteenth century, physicians started to document their findings and actions in personal ledgers [2]. The concept of a single record per patient was introduced in 1907 [3]. A proposal in the 1920s to enforce physicians to document sets of essential data met with much resistance [4]. Since then, the complexity and volume of medical data increased, and specialization led to more health-care workers per patient. As a result, the way data are recorded, processed, retrieved, and communicated became ever more crucial in medical practice [5]. Nevertheless, the patient record underwent very little change [1]. However, it is now generally accepted that the traditional paper patient record can no longer fulfill the expanding demands for information [1]. As an alternative, the computer-based patient record (CPR) increasingly gains interest. The Institute of Medicine (USA) even considers the CPR to be essential for the full maturation of the scientific basis of health care [1].

Storing patient data in an electronic form, already solves some limitations of the paper patient record, namely, availability and illegibility. However, to be more suitable for research, decision support, and quality assessment, it is necessary that data are stored in 'understandable chunks', i.e., structured and coded [6]. Although most CPRs that are in clinical use support coded recording of information to some degree (e.g., diagnoses or medications) [7-11], 'descriptive information' - information traditionally recorded as narratives - often remains recorded as free text [5]. This descriptive information represents an important portion of information on which medical decisions are based. Yet, in an unstructured free-text form, these narratives suffer from incompleteness, ambiguity, and spelling errors [1]. *The research presented in this thesis deals with the problem of obtaining descriptive information in a structured, coded form.*

Obtaining Narratives in a Coded Form

Since Pratt's pioneer work in the 1960s [12], many efforts have been undertaken to extract coded data from free text [13-15]. This technique, referred to as natural language processing (NLP), is attractive, as it does not

interfere with the way in which physicians do their reporting tasks. However, the task is complex, and evidence exists that there is still a long way to go before its implementation in daily practice [13]. Moreover, because the process of data capturing is not influenced, data produced by NLP will still suffer from incompleteness and ambiguity.

The counterpart of NLP are formalisms that are intended to collect narratives in an a priori structured, and coded format, thereby increasing completeness and reducing ambiguity. For many years, paper forms have been used for the collection of structured data for many prospective research purposes. Such paper forms have also proven valuable in clinical care [16]. Their equivalents, referred to as structured data entry (SDE), have been successfully implemented in computer systems [17-22]. However, forms are typically efficient in small, non-complex domains where every 'item' is expected to be filled in. Choosing the same approach for documenting clinical narratives would be impractical, as the domain is usually too large [23]. Most systems that support SDE for larger domains [24-27] are often too limited in their expressive power to cover the complex formulations used in clinical narratives. *Our research involved the development of a general formalism for the a priori capture of structured, coded data with the challenge to approach the expressive power and flexibility of free text.*

Scope

During this research period, a parallel project was conducted, aimed at the development of a general CPR for internal medicine [28, 29]. To gain insight in the requirements of structured data entry we have looked at the domain of internal medicine since, ultimately, SDE of clinical narratives will need to be incorporated in a general CPR. However, to apply and evaluate those insights we chose for endoscopy as a start. The field of endoscopy is large enough to be unsuitable for a form-based approach. At the same time it is sufficiently circumscribed to assess the current status, to develop a prototype, and to conduct an evaluation study.

Endoscopy

Endoscopy is a relatively new discipline in which (parts of) the gastrointestinal tract can be visualized. The procedure can both be

diagnostic and therapeutic. Endoscopists document their findings, performed actions, conclusions and, when considered necessary, therapy and follow-up plan in a report. An endoscopy report has two purposes [30]. Firstly, the report serves as documentation, and can be used as reference material in case the patient has to undergo further endoscopic examinations. Secondly, the report is the primary method to communicate the results of the procedure to the physician who referred the patient for endoscopy.

Analyzing Reporting in Endoscopy

The possibility for any improvement starts with insight in the present situation. Therefore, we started our research *by studying various aspects of current endoscopic reporting*: we studied the extent to which currently produced reports serve the purposes just described, and how endoscopic observations are expressed.

In *Chapter 2* we first describe an inventory to assess the current contents of endoscopy reports. To get insight in the desirability and feasibility of improvement, we asked the opinion of endoscopists on the suitability of reports for follow-up purposes.
In *Chapter 3* we survey the opinions of physicians, who received endoscopy reports upon referral.

In recent years, several steps have been taken to increase the quality of the endoscopic reporting process: a standardized terminology for endoscopy [31] has been proposed, as well as guidelines for the endoscopic investigation and content of reports [32-34]. However, even when such guidelines are used, describing the same topics and using the same terminology does not guarantee that descriptions of the same case by different endoscopists will be identical.

To investigate possible ambiguity in the terminology used, *Chapter 4* describes a study in which endoscopists quantify the terms they use to describe the size of a specific endoscopic lesion.
In *Chapter 5*, we report on the interobserver variability between endoscopists, when they describe morphological features of gastric ulcers, using predefined options.

Supporting the Acquisition of Structured Data

Having gained insight in the content and shortcomings of current free-text reports, we focused on the challenge to capture these data in a structured form while attempting to approach the expressive power and flexibility of free-text reporting.

In *Chapter 6* we describe a general model for structured data entry, based on the notion of explicit descriptional knowledge: knowledge that defines where, how and when medical concepts can be used in clinical narratives. The model allows the construction of specific knowledgebases, each representing the knowledge needed to support SDE within a circumscribed domain. Data capture is made possible through a general entry program, of which the behavior is determined by a combination of user input and the content of the applicable domain knowledgebase. We developed several descriptional knowledgebases, one of which involved the endoscopic domain.

In *Chapter 7*, we describe an evaluation of the model's prototype and the endoscopic knowledgebase. In this evaluation study we aimed to investigate the expressive power of the model, and the influence of formal reporting on the contents of reports. In the *Appendix* of Chapter 7, the complete contents of the gastroscopy knowledgebase is listed.

References

1. Dick RS, and Steen EB, eds. *The Computer-based Patient Record. An Essential Technology for Health Care.* Washington DC: Institute of Medicine, 1991: 15-9.
2. Reiser SJ. The clinical record in medicine. Part 1: Learning from cases. Ann Intern Med 1991; 114: 902-7.
3. Kurland LT, and Molgaard CA. The patient record in epidemiology. Scientific American 1981; 245: 46-55.
4. Reiser SJ. The clinical record in medicine. Part 2: Reforming content and purpose. Ann Intern Med 1991; 114: 980-5.
5. Barnett GO, Jenders JA, and Chueh HC. The computer-based clinical record - Where do we stand? Ann Intern Med 1993; 119: 1042-4.
6. McDonald CJ, and Tierney WM. Computer-stored medical records. Their future role in medical practice. JAMA 1988; 259: 3433-40.

7. Hammond WE, and Stead WW. The evolution of a computerized medical information system. In: Orthner HF, ed. *Tenth SCAMC*. New York: IEEE Computer Society Press, 1986: 147-56.
8. Stead WW, and Hammond WE. Computer-based medical records: the centerpiece of TMR. MD Comput 1988; 5: 48-62.
9. Barnett GO. The application of computer based medical record systems in ambulatory care. New Engl J Med 1984; 310: 1643-50.
10. McDonald CJ, Tierney WM, Overhage JM, et al. The Regenstrief medical record system: 20 years of experience in hospitals, clinics, and neighborhood health centers. MD Comput 1992; 9: 206-15.
11. Van der Lei J, Duisterhout JS, Westerhof HP, et al. The introduction of computer-based patient records in the Netherlands. Ann Intern Med 1993; 119: 1036-41.
12. Pratt AW, and Pacak M. Identification and transformation of terminal morphemes in medical English. Meth Inform Med 1969; 8: 84-90.
13. Baud RH, Rassinoux AM, and Scherrer J-R. Natural language processing and medical records. In: Lun KC et al ed. *Medinfo 92*. Amsterdam: Elsevier Science Publishers, 1992: 1362-7.
14. Haug P, Koehler S, Lau LM, et al. A natural language understanding system combining syntactic and semantic techniques. In: Ozbolt JG, ed. *Eighteenth SCAMC*. Philadelphia: Hanley&Belfus, 1994: 247-51.
15. Sager N, Lynman M, Tick LJ, et al. Natural language processing of asthma discharge summaries for the monitoring of patient care. In: Safran C, ed. *Seventeenth SCAMC*. New York: McGraw-Hill, 1993: 265-8.
16. Wrenn K, Rodewald L, Lumb E, et al. The use of structured, complaint-specific patient encounter forms in the emergency department. Ann Emerg Med 1993; 22: 805-12.
17. McDonald CJ, Tierney W, and Blevins L. The benefits of automated medical record systems for ambulatory care. In: Orthner HF, ed. *Tenth SCAMC*. New York: IEEE Computer Society Press, 1986: 67-83.
18. Kent DL, Shortliffe EH, Carlson RW, et al. Improvements in data collection through physician use of a computer-based chemotherapy treatment consultant. J Clin Oncol 1985; 3: 1409-17.
19. Lussier YA, Maksud M, Desruisseaux B, et al. PureMD: a computerized patient record software for direct data entry by physicians using a keyboard-free pen-based portable computer. In: Frisse ME, ed. *Sixteenth SCAMC*. New York: McGraw-Hill, 1992: 261-4.

20. Kuhn K, Gaus W, Wechsler JG, Janowitz P, et al. Structured reporting of medical findings: Evaluation of a system in gastroenterology. Meth Inform Med 1992; 31: 268-74.
21. Bell DS, and Greenes RA. Evaluation of UltraSTAR: Performance of a collaborative structured data entry system. *Eighteenth SCAMC.* New York: McGraw-Hill, 1994: 216-22.
22. Gouveia-Oliveira A, Raposo VD, Salgado NC, et al. Longitudinal comparative study on the influence of computers on reporting of clinical data. Endoscopy 1991; 23: 334-7.
23. Van Ginneken AM, Stam H, and Moorman PW. A multi-strategy approach for medical records of specialists. In: Greenes RA et al, eds. *Medinfo 95.* Edmonton: HC & CC, 1995: 290-4.
24. Nowlan WA, and Rector AL. Medical knowledge representation and predictive data entry. In: Fieschi M, Talmon J, eds. *Proceedings of AIME-91.* Heidelberg: Springer Verlag, 1991: 105-16.
25. Trace D, Naeymi-Rad F, Haines D, et al. Intelligent medical record - entry (IMR-E) J Med Sys 1993; 17: 139-51.
26. Campbell KE, and Musen MA. Creation of a systematic domain for medical care: the need for a comprehensive patient-description vocabulary. In: Lun KC et al ed. *Medinfo 92.* Amsterdam: Elsevier Science Publishers, 1992: 1437-42.
27. Campbell KE, Wieckert K, Fagan LM, and Musen MA. A computer-based tool for generation of progress notes. In: Safran C, ed. *Seventeenth SCAMC.* New York: McGraw-Hill, 1993: 284-8.
28. Van Ginneken AM, Van der Lei J, and Moorman PW. Towards unambiguous representation of patient data. In: Frisse ME, ed.*Sixteenth SCAMC.* New York: McGraw-Hill, 1992: 69-73.
29. Van Ginneken AM, Stam H, and Duisterhout JS. A powerful macro-model for the computer patient record. In: Ozbolt JG, ed. *Eighteenth SCAMC.* Philadelphia PA: Hanley&Belfus, 1994: 496-500.
30. Pfeil SA. The upper gastrointestinal endoscopy report. Gastrointest Endosc Clin 1994; 4: 541-50.
31. Maratka Z. *Terminology, definitions and diagnostic criteria in digestive endoscopy, 2nd ed. OMED database of digestive endoscopy.* Bad Homburg: Normed Verlag, 1989.
32. Quality assurance of gastrointestinal endoscopy. An information resource manual from the American Society for Gastrointestinal Endoscopy. Manchester MA, 1989.

33. Appropriate use of gastrointestinal endoscopy. American Society for Gastrointestinal Endoscopy. Manchester MA, 1992.
34. Crespi M, Delvaux M, Schapiro M, et al. Minimal standards for a computerized endoscopic database. Am J Gastroenterol 1994; 89 Suppl: 144-53.

Chapter 2

The contents of free text endoscopy reports: An inventory and evaluation by peers

Moorman PW [a,b]
van Ginneken AM [a]
van der Lei J [a]
Siersema PD [b]
van Blankenstein M [b]
and Wilson JHP [b]

[a] Department of Medical Informatics, Erasmus University,
[b] Department of Internal Medicine, University Hospital Dijkzigt,
Rotterdam, The Netherlands

Endoscopy 1994; 26: 531-8

Abstract

Insight in the current status of endoscopy reports is needed for a discussion on the desirability and feasibility of (more) standardized endoscopy reporting.

We collected, from 10 endoscopists, 181 reports in two diagnostic and two indication categories. An inventory was made of subjects dealt with in the reports such as: indication, premedication, therapy plan and descriptive aspects of ventricular ulcers and lower tract polyps. To assess endoscopists' opinions on their reports, 16 randomly selected reports were reviewed by the 10 endoscopists, according to the Delphi method.

The reports varied enormously in content and detail; 19 of the 28 subjects were not explicitly described in more than 50% of the studied reports. Such variation in the contents of reports may decrease quality of care. The large number of topics that endoscopists indicate to be missing in their reports (on average 14 topics per report) suggests that more detail should be given in endoscopy reports. The current method of reporting causes endoscopists to omit information, that they consider important. Due to the low overall consensus among endoscopists on inclusion of specific topics (on only 15% of all topics 8 or more endoscopists agreed), we conclude that general criteria for the contents of reports cannot yet be formulated. However, the fact that the endoscopists agreed with more than one third of remarks made by colleagues, opens perspective to identify criteria for the formalization of certain report categories.

Introduction

Endoscopy has become an important diagnostic and therapeutic procedure in clinical medicine. The procedures and findings of endoscopy examinations are documented in clinical reports. As important decisions are based on the information in reports, the contents need to be reliable, complete and unambiguous. This can only be achieved when not only an interpretation of the findings (e.g. gastritis) is given in a report, but also a description independent of the interpretation (e.g. red and swollen mucosa).

Kuhn et al [1] concluded from a multi-center study that descriptions of findings were insufficient in a large percentage of cases: omissions ranged from 4.3% for gastric ulcer size to 12.5% for grading of oesophageal varices.

To address the limitations of paper-based free text reports, the possibilities of standardized terminology [2], guidelines for quality assurance [3], and of computer systems assisting in report management [4-8] are being explored. Gouveia-Oliveira et al [9] and Mai et al [10] showed that reminders, either in the form of peer review or structured data entry, increase the information content of reports. Other researchers, however, have shown that such positive effects may only be temporary [11].

The dilemma of describing observations in free text versus structured data entry can be formulated as the conflict between individualism and uniformity. In individualism, each endoscopist records what he or she considers to be important. Uniformity, however, is a prerequisite for standardization and reduction of ambiguity. It demands from endoscopists to use the same terminology, and to describe the same subjects and features under comparable circumstances. As terminology has been the subject of most discussions on uniformity, we focus our research on which information endoscopy reports should contain.

We believe that, for the benefit of uniformity, it is justified to encourage endoscopists to describe more than they are used to, but they should never be limited in describing what they consider important. An optimal coded data set, within practical limits [12,13], requires insight in the overlap or variety of current reports and the motivation of endoscopists to conform with peers. To investigate what degree of uniformity in endoscopy reporting already exists, and what degree is feasible, we conducted two related studies.

The first study entailed an inventory of the contents of endoscopy

reports, in order to establish:

- whether reports of different endoscopists contain similar information on comparable subjects.

To obtain insight in the opinions of endoscopists a second study, according to the Delphi method [14, 15], was performed. This method not only provides insight in current consensus, but also enhances consensus. The Delphi study would enable to assess:

- whether endoscopists feel that endoscopy reports contain sufficient information,
- whether the opinions of endoscopists, with regard to what should be recorded, are fundamentally different or overlapping,
- on which subjects a consensus can be reached that allows defining minimal data sets (MDS).

Methods and Materials

We asked ten endoscopists (one from the Dijkzigt University Hospital in Rotterdam, and one from each of the nine affiliated teaching hospitals in the Rotterdam region) to participate in the two studies. All endoscopists agreed to participate. They had received their endoscopy training at different Dutch teaching hospitals.

Report collection

From each endoscopist, endoscopy reports were scanned as filed; reports that met the following criteria were included in the study:

- the described endoscopy examination had to be a first endoscopy for the patient,
- the report had to have been made before the endoscopists were asked to participate in the study (July first 1991),
- the report had to fall in one or more of the following 4 categories: (1) the endoscopic diagnosis is gastric ulcer (= ulcer category), (2) the endoscopic diagnosis is polyp in the lower gastro-intestinal tract (= polyp category), (3) the indication for the endoscopy examination is vague abdominal complaints (= dyspepsia category), and (4) the indication is bleeding (= bleeding category),
- per endoscopist and per category the first 5 retrieved reports.

Instead of 200, we collected 181 reports: 7 reports fell into more than one category, 1 polyp report proved to be a follow-up examination, and from one endoscopy unit only 4 ulcer category reports and no indication category reports could be obtained, because the used computer system had only been recently introduced, and indications were not included in the reports.

Thus, 49 reports were collected in the ulcer category, 49 in the polyp category, and 45 in both indication categories.

Analysis of inventory

We divided the identified subjects in the reports into (1) subjects denoting features of ulcers and polyps (such as ulcer base and polyp size) and (2) subjects that are more or less findings-independent and appeared regularly throughout the reports (e.g. indication and premedication). These latter subjects are referred to as 'general parameters'.

First, we prepared a list of all general parameters and ulcer- and polyp features, that occurred in the reports. Subsequently, the subjects were counted in the appropriate reports. The identified general parameters - *indication, therapeutic plan,* and *follow-up* - were only counted in the diagnosis categories, whereas *premedication, type of used endoscope,* and information on the *extent of the examination* were counted in all reports. (For reports in the indication categories indication was the selection criterion. In many of the examinations described in the indication category reports, no or only minor abnormalities were found. In such cases, although not explicitly stated, therapy or follow-up is not considered relevant.) The general parameter *location of the squamocolumnar mucosal junction,* or Z-line, was only counted in oesophago-gastro-duodenoscopy reports.

In 7 reports in the polyp category more than 3 polyps were encountered. Because the relatively large number of polyps contributed by one report would introduce a bias, these 7 reports were omitted in the analysis of polyp descriptions.

Descriptions of subjects that did not permit unambiguous interpretation, were considered insufficient. With respect to the 'extent of an endoscopy examination', for example, explicit descriptions (in cm's or anatomically), and implicit statements (such as 'the descending part of the duodenum is normal') were considered sufficient, whereas statements such as 'duodenum: normal' were considered insufficient.

Delphi study

After a short pilot study to assess the time required to evaluate a report, we randomly selected 16 reports from the collected 181 reports. The 16 reports consisted of 4 reports per category; at least one, and no more than 2 reports from each endoscopist were included. All selected reports were re-typed in a standardized format. Any information that could lead to the identification of hospital, endoscopist or patient was omitted.

In the first round, all 16 reports, in different random order for each endoscopist, were sent to the 10 endoscopists with the request to indicate for each report whether the information was sufficient for proper patient follow-up. If they felt information was missing in a report, they were asked to describe what information should have been included.

We thus obtained 160 evaluations of reports.

To compare the different evaluations of an individual report, all remarks, this is any information reported to be missing, were analyzed. Remarks relating to the same report and conveying the same meaning were mapped to one phrase representing that meaning. For example, if one endoscopist had written 'is the base of the ulcer regular?' and another had remarked 'the base of the ulcer is not described', then these remarks were mapped to the standard phrase 'In this report, the base of the ulcer should have been described'. The proposed mapping was checked and corrected by an independent jury of two expert endoscopists.

In the second round, all reports were again sent to the participating endoscopists. Each report was now accompanied by all remarks made on that report, in the transcribed form. For each remark, the number of endoscopists that had made the remark spontaneously in the first round of the study, was indicated. The endoscopists were asked to state whether or not they agreed with a remark.

Results

Results of the inventory

General parameters (Table 1). Except for indications, all general parameters were present in less than half of the reports. Furthermore, there was considerable variation in the way these parameters were described. Indications were either formulated as a complaint of the patient, as a finding of an other examination than endoscopy, e.g. polyp found on X-ray examination, or as a statement of which the origin could not be derived, e.g.

suspicion of polyp. In one third of the 21 reports in which medication was mentioned as the therapeutic plan, it was unclear whether the medication was advised or prescribed. Only half of the given follow-up advice stated when the follow-up examination had to be performed. In none of the reports was explicitly stated that follow-up of the patient was not necessary. In the 19 reports where premedication was recorded as being given, 4 times the name of the drug, the dosage and administration route were described, whereas in the remaining reports only the name of the drug was mentioned.

Table 1. Frequency of general parameters.

Subject	Number of reports	Number of times mentioned
Indication	98[a]	
Absent		24
Present		74
Therapeutic plan	98[a]	
Absent		74
Operation		2
Medication		22
Follow-up	98[a]	
Absent		50
Present		48
Premedication	181[b]	
Absent		152
Not given		10
Given		19
Type of used scope	181[b]	
Absent		102
Present		79
Extent of examination	181[b]	
Unclear		59
Explicit anatomical		61
Explicit in cm's		14
Explicit combination		5
Implicit anatomical		42
Location Z-line	105[c]	
Absent		48
Distance to incisors and relative to diaphragm		16
Distance to incisors or relative to diaphragm		41

[a] Reports in ulcer and polyp categories.
[b] All reports.
[c] Gastroscopy reports.

Descriptions of ulcers and polyps. In the 49 reports collected in the ulcer category, we encountered 59 descriptions of ulcers. In the 42 'lower tract polyp' reports, 54 polyp descriptions were found. Described ulcer and polyp features are listed in Tables 2 and 3, respectively, together with the

frequency in which the features appeared.

Table 2. Described ulcer features, found in 59 ventricular ulcer descriptions, encountered in 49 reports.

Feature	Number of times mentioned
Localization	58
Size	40
Depth	17
Shape	7
Base	14
Border	11
Surroundings	7
'Punched out'	10
Bleeding stigmata *	13
Biopsy information	43
Interpretation **	26

* Absence or presence of bleeding, clots, or visible vessels

** Such as malignant, benign or suspect

Table 3. Described polyp features, found in 54 polyp descriptions, encountered in 42 reports.

Feature	Number of times mentioned
Localization	54
Size	36
Diameter	18
Shape	9
Sessile/ Pedunculate	32
Color	6
Surface	4
Bleeding stigmata	4
Biopsy / Polypectomy	54
Interpretation	16

Even when a given feature is mentioned in a report, the feature may not be described completely. Descriptions of an ulcer base include, among others, the presence or absence of exudate, colour or type of exudate and regularity of the base. The most extensive base descriptions included statements about the colour and regularity, and were encountered in 2 of the 14 base descriptions.

A numerical description of the size was given in half of the cases that an ulcer size was described, and in one fifth of the polyp size descriptions. In the non-numerical ulcer size descriptions, we encountered 11 different terms, ranging from small to gigantic.

With regard to information about biopsies taken from ulcers: 5 reports explicitly stated that a biopsy had not been performed, whereas the site (border or base) and the number of biopsies were mentioned in very few reports.

All polyp descriptions stated explicitly whether a biopsy or polypectomy had, had not been, or was going to be performed. Of the 31 cases in which a polypectomy had been performed, additional information on the method of removal was given in 23, macroscopic partial or complete removal was mentioned in 5, and the presence or absence of bleeding after removal was stated in 4 cases.

None of the ulcer and polyp descriptions contained information about all features listed in Tables 2 and 3. For ulcers, at most 8 of the 11 features

were described (3 of 59 ulcer descriptions), whereas in the most extensive polyp description 7 of the 10 features were described (1 of 54 polyp descriptions).

Results of Delphi

In the second study, 10 endoscopists reviewed the 16 selected reports in two rounds. In the first round the participants indicated missing information. In the second round they gave their opinion on the made remarks.

General results of the first and second round. In the first round, for each of the 16 reports, the number of endoscopists that judged a report to contain insufficient information ranged from 3 to 10, with an average of 7.6. The endoscopists together made on average 29 remarks on a report. After mapping to transcriptions conveying the meaning of the original remarks, we counted on average 14.25 different remarks per report (range 7-25). As the average number of different remarks is smaller than the number of total remarks, some remarks were made by more than one endoscopist. The average number of endoscopists that made a particular remark was 2.0 (29/14.25).

In the second round, the number of endoscopists judging a report to contain insufficient information ranged from 8 to 10 (average 9.6). The number of reports considered to be incomplete by all 10 endoscopists increased from 2 to 11. On average, the endoscopists agreed on more than twice as many remarks (65) than made in the first round. The average number of endoscopists supporting a particular remark increased from 2.0 to 4.5.

The remarks. Since, on average, 14.25 different remarks were made on each report, in total 228 different remarks were made. The term *different* relates to remarks made on different reports. Thus, when a remark is made on two reports, it is counted twice. The remarks covered a wide spectrum, and can be divided into three broad categories: remarks concerning a missing general part, concerning uncertainty of what had been examined, and concerning an insufficient description of an observed abnormality.

About 35% of the 228 different remarks related to an absent general part of a report, 23% to uncertainty of what had been examined: uncertainty

about the extent of the examination, the presence or absence of abnormalities, and the reliability of observations. The subjects of these remarks are listed in Table 4.

Table 4. Remarks made on missing general information and uncertainty about examination.

Subject	Reports with remark	Average number of agreeing endoscopists	
		Round 1	Round 2 (range)
Indication	3	4.7	8.0 (8-8)
Premedication	11	1.6	3.6 (3-6)
Type of scope	9	3.0	5.3 (4-6)
Therapy plan	5	1.6	3.2 (3-4)
Follow-up	3	1.3	4.0 (2-7)
Location Z-line	8	2.4	7.1 (6-8)
Extent of gastroscopy	5	1.8	4.8 (4-5)
Extent of sigmoidoscopy	3	5.6	7.0 (4-9)
Absense or presence of hiatus hernia	3	1.3	6.7 (6-7)
Absence or presence of blood	2	1.0	7.0 (7-7)
Answer to indication	1	4.0	9.0 (-)
Reliability of examination	6	1.2	4.5 (4-5)

To give an example, the remark 'The location of the Z-line should have been described in this report' was made on 8 reports. In the first round, on average 2.4 endoscopists made this remark, while in the second round 6 - 8 endoscopists (on average 7.1) agreed with the remark.

Although most remarks received more support in the second than in the first round, support for some remarks increased only slightly; remarks about premedication and therapy plan were supported by 2 more endoscopists in the second than in the first round.

While some endoscopists always gave the same response to a specific remark in different reports, others agreed with a remark in some and disagreed with the same remark in other reports. The remark stating that the type of the scope used for the examination should be included in the report, was made on 9 reports: 3 endoscopists agreed with the remark all 9 times, 3 disagreed all times and 4 varied.

Sometimes the number of agreeing endoscopists varied with the type of examination. Remarks about absent premedication information were on average supported by 3 endoscopists for gastroscopies and 6 for colonoscopies, whereas no such remarks were made on sigmoidoscopy reports lacking premedication information.

About 42% of the 228 different remarks related to insufficient descriptions of abnormalities. As the selection criterion for 8 of the 16 reports was either ventricular ulcer or lower tract polyp, most remarks were made on incomplete descriptions of ulcers and polyps. In Table 5, *all* remarks are listed that were made on absent descriptions of ulcer and polyp features. With respect to ulcers, remarks about an absent description of the numerical size, the border and bleeding stigmata received most support from the endoscopists after the second round. For polyps, a numerical size description, characterization as sessile or pedunculate and the method of removal were considered to be most important.

Table 5. Remarks made on ulcer and polyp descriptions.

Subject	Descriptions with remark	Average number of agreeing endoscopists	
		Round 1	Round 2 (range)
Ulcer			
Numerical size	5	5.4	9.0 (8-10)
Depth	4	1.8	4.8 (4-6)
Shape	4	2.0	4.5 (3-5)
Base	6	3.0	5.8 (4-7)[a]
Border	5	3.2	6.8 (3-9)[a]
Surroundings	5	1.6	2.6 (1-5)
Bleeding stigmata	4	4.5	7.0 (6-8)
Interpretation	2	2.5	7.0 (7-7)
Polyp			
Numerical size	5	2.8	8.2 (5-9)[b]
Sessile/pedunculate	3	2.0	7.7 (6-9)[b]
Color	6	1.0	1.5 (1-2)
Surface	6	1.2	3.5 (2-5)
Shape	1	1.0	5.0 (-)
Interpretation	1	1.0	4.0 (-)
Method of removal	2	3.0	8.0 (8-8)
Complete/partial removal	1	2.0	6.0 (-)

[a] The lowest number of agreeing endoscopists was encountered in an ulcer described as being covered by a blood clot.

[b] The lowest number of agreeing endoscopists was encountered in a polyp described as 'giant' and 'occluding the lumen'.

Endoscopist variation and consensus. As stated above, on average an endoscopist made 2.9 remarks per report in the first round and supported 6.5 remarks in the second round. There were large differences between the endoscopists with respect to these averages: they ranged from 0.2 - 6.25 remarks per report in the first, and 3.0 - 8.9 in the second round.

In the second round, the endoscopists agreed on average with 83% (range 71%-100%) of the remarks made by themselves in the first round. When confronted with a new remark, thus a remark made by a colleague, and not

by themselves, the endoscopists agreed on average on 34% (range 14%-51%) of these remarks.

In Figure 1 the 228 different remarks are stratified per number of endoscopists that made and/or agreed with a remark. In the first round, the highest agreement was 7 endoscopists; this was the case for 3 of the 228 remarks. Highest agreement in the second round was 10, again for 3 remarks. Two of these remarks concerned an absent numerical ulcer size, the third concerned an absent description of a found gastritis.

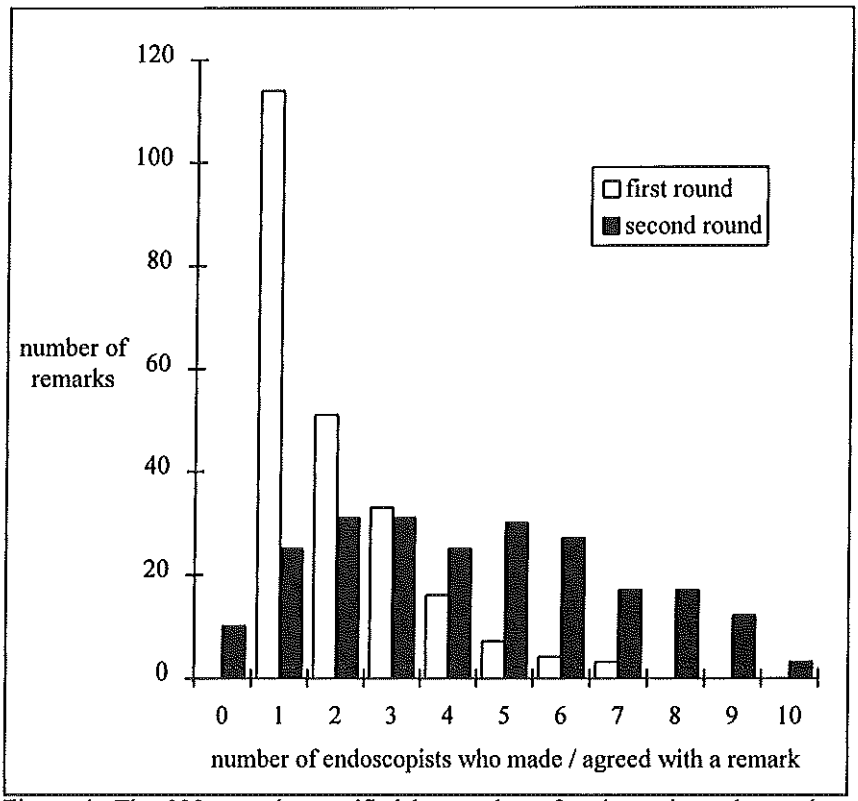


Figure 1. The 228 remarks stratified by number of endoscopists who made a remark in the first round, and agreed with a remark in the second round of the Delphi study.

The vast majority, 186 remarks, got more support in the second than in the first round. Sixteen remarks received less support in the second than in the first round, of which 10 remarks, all initially made by one endoscopist, were no longer supported by any endoscopist. Of the 26 remarks where an

equal number of endoscopists agreed with that remark in the first and second round, only 15 got the same support from the same endoscopists in both rounds.

Evaluation of own reports. In the first round 7 evaluations of endoscopists were without remarks on their own reports, in the second round this decreased to 2. In the first round the endoscopists made on average 1.1 remarks on their own reports. After they were confronted with the remarks of their colleagues this average increased to 4.3.

Discussion

Inventory

In the first study we made a detailed inventory of the contents of 181 free text endoscopy reports, to obtain insight in the current degree of uniformity of endoscopy reports.

We found large variation in the contents of endoscopy reports. The endoscopists recorded different general parameters, in a non-uniform way. It is striking that not a single general parameter was recorded in all reports. The ulcer and polyp descriptions contained, in general, little detail and, in different reports, different features were described.

Variation in reports may lead to ambiguity. Clinical interpretation of a description may be difficult when a large set of overlapping terms are available to describe a feature. It may be expected that there is limited consensus on the meaning of 11 different verbal expressions of an ulcer size [16]. Difficulty in interpretation may also occur when a report does not state whether or not a biopsy has been performed. A general rule, dictating when, and when not to include biopsy information, would make the interpretation of an absent biopsy statement self-evident. Such general rules are not in use: some ulcer reports contain explicit statements that a biopsy was not performed, while other reports do not contain any information about a biopsy. The same type of ambiguity exists when it is not clear whether medication is only advised or also prescribed, which may lead to double medication or none at all. Not only ambiguity, but also the simple absence of information may decrease quality of care. When information is not recorded, data that can serve as a reference value will not be available e.g. to assess a change in ulcer size at a follow-up examination.

Kuhn [1] and Mai [10] report frequencies of omissive errors for specific subjects. Although our study was intended to assess variability, and not omissive errors, and our study is based on a smaller number of reports, it suggests higher omissive frequencies of equal subjects (e.g. ulcer size). Comparison of the results remains difficult as (a) there may be differences between reports from academic and non-academic hospitals, (b) ambiguous statements, in our study treated as absent descriptions, may account for differing absence frequencies. The fact that none of the subjects was described in all reports, suggests that far more types of omissive errors exist than those studied by Kuhn and Mai.

If we assume that what endoscopists write in their reports is important, then it is strange to find a large variation in contents. Do endoscopists differ with regard to what they consider important information in reports, or is the variation due to the fact that they do not bring to practice what they think is important? Why could intuitive rules not be found, such as 'in case of diagnosis A, feature B must always be described' or 'a report should give an answer to the clinical question' (the absence or presence of a source for the blood loss was only given in 19 of the 45 reports with an indication bleeding)?

More uniform and more complete reports will require endoscopists to record more information than they currently do. Such a change in reporting behaviour will only be feasible if endoscopists themselves consider such a change necessary.

Delphi study

To assess whether endoscopists indicate that more complete and uniform reporting is considered desirable, we have performed the second study.

Endoscopists indicate that more detail should be given in their reports. This is supported by the fact that none of the reports was considered to contain sufficient information and, in the end, every report received comments from at least 8 endoscopists. Secondly, endoscopists comment and agree on remarks with respect to their own reports. Furthermore, the fraction of endoscopists agreeing on the necessity to describe a subject is (much) larger than the fraction of reports in which that subject was actually mentioned (Table 6). For example, although on average 90% of the endoscopists agree that the ulcer size should be described numerically, this had only been done in 32% of the ulcer descriptions.

Table 6. Actual presence of subjects in reports versus desirability of subjects.

Subject	Present in % of applicable cases (N)*	Mean fraction of agreeing endoscopists on remark about absence (N)**
Ulcer feature		
Numerical size	32% (59)	90% (5)
Base	24% (59)	58% (6)
Border	19% (59)	68% (5)
Bleeding stigmata	22% (59)	70% (4)
Polyp feature		
Numerical size	13% (54)	82% (5)
Sessile / Pedunculate	59% (54)	77% (3)
Surface	7% (54)	35% (6)
Polypectomy method	74% (32)	80% (2)

* N refers to the number of descriptions in which the presence of the feature was counted.

** N refers to the number of cases in which a description of the feature was absent, and a remark was made on the feature's absence.

We also conclude from Table 6 that the contents of endoscopy reports are not a reflection of the opinions of the endoscopists. The variability in endoscopy reports thus stems, for a large part, from a discrepancy between what endoscopists think should be included in a report and what they actually record.

Although all endoscopists indicate a need for more information in their reports, this study does not answer *which* information should be provided, as there is little overlap in what the individual endoscopists indicate to be necessary additions. Firstly, this is illustrated by the, on average, low number (4.5) of endoscopists agreeing on inclusion of a subject in a report. Note, that in such cases, about half of the endoscopists support and about half of the endoscopists do not support inclusion of a subject in a report. Secondly, *all* 10 endoscopists agreed on only 3 of the 218 remarks. Thirdly, when agreement of 8 or more endoscopists would be considered acceptable consensus, then consensus could only be reached on 15% of all subjects. Based on 7 or more endoscopists this would be 22%.

Because of this poor consensus it is not possible to define, on basis of this study, a minimal set of subjects by which an examination or finding should be described. Even when the requirements for the contents of endoscopy reports would be based on a cross-section of the opinions of 7 or more endoscopists, reports meeting those requirements would contain so little information that they would elicit even more remarks than is presently the case.

This does however not imply that structured data entry, or the formulation of Minimal Data Sets is impossible, as the fractions of endoscopists agreeing on inclusion of a subject do not necessarily reflect the highest achievable support:

- Disagreement with a remark does not directly imply that such information is considered irrelevant. An endoscopist may also interpret a report in such a way, that he considers the information already implicitly present in the report. Such interpretations may be strongly influenced by the endoscopists own habits. If an endoscopist only mentions the preparation of the bowels when reliable judgement of the mucosa is not possible, he will assume that observations are reliable unless explicitly mentioned otherwise.
- The opinions of the endoscopists do not seem to be based on *strong* convictions about what should be included in a report. Remarks were not always made on all reports lacking the same information, and responses to remarks appearing in several reports varied. The statement 'duodenum normal' elicited only in 5 of 7 reports containing that statement, the remark 'which parts of the duodenum have been examined ?', while 4 endoscopists constantly changed their opinion on inclusion of the type of used endoscope. Some inconsistencies may be explained by differences in the information content of the report as a whole. The observation that the remark 'the ulcer border must be described' received 4 times agreement from about 8 endoscopists, and in one case only from 3, may be explained by the fact that this specific ulcer was covered by a blood clot.
- The endoscopists were quite accepting of the comments of their colleagues. On average, the endoscopists supported more than one third of the remarks that were brought to their attention by colleagues in the second round.

Based on gained insights we can, however, make some suggestions for minimal data sets. If data sets are going to be used for report making, then this should not be done on the philosophy that every subject in that set should always be described. There should be several sets, applicable to reports having parameters in common such as type of examination, findings, and indication. Given these parameters, it may be feasible to identify at least two types of subjects. 1) Subjects that should always be described unless the reason is explicitly stated that a description is not

possible. For example, stomach mucosa is always present, but it may have been difficult to examine because of the presence of blood. 2) Subjects, whose absence or presence should be explicitly mentioned, depending on the context of the report. E.g. if the indication is 'suspicion of varices', then the presence or absence of varices should be explicitly mentioned.

Conclusions and Recommendations

A large variation exists in the content of endoscopy reports. Our Delphi study shows that endoscopists indicate that more uniform and more complete reports are desirable and feasible. Low overall consensus did not permit us to define which information should at least be present to improve uniformity.

Besides the ongoing discussion on endoscopic terminology, a discussion is also needed on which information needs to be recorded in endoscopy reports, criteria when this information needs to be recorded and how this information can be made explicit.

Meanwhile, the effect of methods other than free text reporting should be explored to reduce the discrepancy between what should be recorded and what actually is recorded. Using the information from this study, we are presently developing strategies to combine reduction of ambiguity and omissions with efficient structured data entry.

Acknowledgement

We thank the endoscopists for participating, making their reports available for this study, and helping in mapping the remarks. This study was financially supported by Glaxo BV.

References

1. Kuhn K, Swobodnik R, Johannes RS, et al. The quality of gastroenterological reports based on free text dictation: An evaluation in endoscopy and ultrasonography. Endoscopy 1991; 23: 262-4.
2. Maratka Z. *Terminology, definitions and diagnostic criteria in digestive endoscopy, 2nd ed. OMED database of digestive endoscopy.* Bad Homburg: Normed Verlag, 1989.
3. Quality assurance of gastrointestinal endoscopy. An information

resource manual from the American Society for Gastrointestinal Endoscopy, 1989.
4. Stolzing H, Birkner B, Lindlar R, et al. Computerunterstutzte dokumentation bei der oberen gastrointestinalen endoscopie: Erfahrungen bei der routineanwendung in drei kliniken. Z Gastroenterol 1989; 27: 667-75.
5. Kruss DM. The ASGE database: computers in the endoscopy unit. Endoscopy Rev 1987; 4: 64.
6. Nguyen HN, Walker S, and Bode JC. Routineeinsatz eines computer-unterstutzten dokumentationssystems fur die osophago-gastro-duodenoskopie. Z Gatroenterol 1991; 29: 121-4.
7. Aabaken L, Westerheim J, Hofstad B, et al. SADE database for endoscopic procedures: aspects of clinical use. Endoscopy 1991; 23: 269-71.
8. Gouveia-Oliveira A, Raposo VD, Azevedo AP, et al. SISCOPE: a multiuser information system for gastrointestinal endoscopy. Endoscopy 1991; 23: 272-7.
9. Gouveia-Oliveira A, Raposo VD, Salgado NC, et al. Longitudinal comparative study on the influence of computers on reporting of clinical data. Endoscopy 1991; 23: 334-7.
10. Mai HD, Sanowski RA, and Waring JP. Improved patient care using the A/S/G/E guidelines on quality assurance: a prospective study. Gastrointest Endosc 1991; 37: 597-9.
11. Lomas J, Anderson GM, Domnick-Pierre K, et al: Do practice guidelines guide practice? The effect of a consensus statement on the practice of physicians. New Engl J Med 1989; 312: 1311-5.
12. De Dombal. Organization of data input- the importance of rapid/high quality data collection. Endoscopy 1992; 24 suppl 2: 490-2.
13. Sobala GM, and Axon ATR. Data processing in endoscopy. Endoscopy 1992; 24: 167-8.
14. Dalkey NC. An experimental study of group opinion: The Delphi method. Futures 1969; 1: 408-26.
15. Romm FJ, and Hulka BS. Developing criteria for quality of care assessment: Effect of the Delphi technique. Health Serv Res 1979; 14: 309-12.
16. Toogood JH. What do we mean by 'usually'? Lancet 1980; 17: 1094.

Chapter 3

Are referring physicians satisfied with endoscopy reports?

Moorman PW [a,b]
Siersema PD [b]
van Ginneken AM [a]
and van Blankenstein M [b]

[a] Department of Medical Informatics, Erasmus University,
[b] Department of Internal Medicine, University Hospital Dijkzigt,
Rotterdam, The Netherlands

Z Gastroenterol 1994; 32: 623-5

Abstract

To assess the opinions of referring physicians on the contents of endoscopy reports, 150 consecutive endoscopy reports were accompanied by a questionnaire. Of these, 102 reports were returned: response was 68%. Almost half of the reports were considered not fully satisfactory. However, endoscopy reports may be improved by including information such as indication, therapy plan and follow-up plan on a more regular basis, and add clarity whether findings may account for complaints of the patient.

To tailor endoscopy reports to the needs of individual referring physicians, more explicit information of referring physicians is required. If endoscopists are responsible for the information they provide to the referrer, it is also their task to facilitate the explicit formulation of preferences by the referrer.

Introduction

Endoscopists document the findings of their examinations in endoscopy reports; these reports subsequently play an important role in communicating the results of the examination to the referring physician. Despite this important role, little is known about the referring physician's opinion regarding the content and quality of endoscopy reports.

Research in other areas of medicine, however, has shown inadequacies in reporting. In a survey of radiology reports, 40% of the physicians indicated that they found radiology reports occasionally confusing [1]. In a study of discharge reports from dermatology, gastroenterology, neurology and internal medicine departments, a panel of general practitioners and specialists judged 21% of the reports as poor or barely adequate [2].

An endoscopy report is also used by endoscopists as reference material for follow-up examinations. In a previous study, endoscopists indicated that 96% of the evaluated reports contained insufficient information for follow-up purposes [3].

Recently, investigators have argued that computerisation may improve the quality of endoscopy reporting [4-9]. The ASGE guidelines, for example, promote inclusion of a therapy and follow-up plan [7]. Computerisation of endoscopy reports would also enable an automatic tailoring of reports to the needs of the receiver of the report; a referring physician could receive a report more concise than the report to be used as reference for follow-up examinations by the endoscopists themselves. And even different reports could automatically be created for different referring physicians, if their information requirements would turn out to be different.

As computerisation of endoscopic reporting may have impact on the content of reports, we performed a survey among physicians who referred patients to our endoscopy unit to assess:

- their opinion on the quality of current endoscopy reports,
- possible differences in the information preferences of the various referring physicians,
- how the contents of endoscopy reports may be improved,
- possible implications for computerised reporting.

Methods and Materials

Starting March 1993, we enclosed a questionnaire with 100 consecutive

endoscopy reports to the referring physician working in internal medicine (internist) and with 50 consecutive reports to the referring general practitioner (GP). We limited the study to referring physicians who did not perform endoscopies themselves. The reports were produced at the gastroenterology unit of the University Hospital Dijkzigt Rotterdam, the Netherlands; the same hospital where the internists work. In our unit, endoscopy reports are dictated and typed out on a form separate from the form on which the endoscopy was requested.

Each GP is free to refer a patient to the endoscopy unit in our or to another hospital in the same area. Formal guidelines, about when to refer a patient to an endoscopy unit, or the content of endoscopy reports, are not in use. Nor are there any special joint educational activities aimed at improving communication.

Table 1. Questions and answers on the questionnaire.

Questions	Possible answers
Reason for referral	Open-ended
Is an indication present?	Yes / No
if no: should it have been included?	Yes / No / No opinion
The findings description is	Too detailed / Sufficiently detailed / Insufficiently detailed / No opinion
Does this report answer your clinical question?	Yes / No
if no: which question remained unanswered?	Open-ended
Is it sufficiently clear whether findings can account for the complaints?	Yes / No
Is a therapy plan included?	Yes / No
if yes: do you agree with the therapy plan	Yes / No
if no: do you believe therapy is required?	Probably not necessary / Probably necessary, and known which / Not known whether or which therapy is necessary
Would you have wanted a therapy advise?	Yes / No / No opinion
Is a follow-up included?	Yes / No
if no: do you believe follow-up is required?	Probably not necessary / Probably necessary, and known when / Not known whether or when follow-up is necessary
Would you have wanted a follow-up advise?	Yes / No / No opinion

Table 1 shows the content of the questionnaire. Questions pertained only to the report to which it was attached.

A report was considered not to be completely satisfactory to the referring physician when he/she indicated that one or more of the following inadequacies were present in the report:

- an indication was absent, but desired,

- findings were described in too much or insufficient detail,
- it was insufficiently clear whether or not the described findings could account for the complaints of the patient,
- the clinical question was insufficiently answered,
- a therapy or follow-up plan was absent, but desired,
- a therapy or follow-up plan was present, but undesired.

Results

Of the 150 questionnaires 102 were returned: overall response was 68% (internists 69%, GPs 66%). The group of responding internists consisted of 24 different internists, of whom 3 returned up to 5 reports. All questionnaires returned by GPs came from 33 different GPs. General information about the 102 study reports is listed in Table 2.

Table 2. General information on the 102 study reports.

Subject	Occurrences
Referring physician	
Internist	69
General practitioner	33
Type of examination	
Gastroscopy	70
Sigmoidoscopy	21
Colonoscopy	11
Clinical question	
Abdominal complaints	60
Bleeding	13
Follow-up of sclerotherapy, polypectomy or gastric ulcer	11
Suspicion of polyp or carcinoma lower tract	9
Other	9
Indication absent in report	46
Therapy plan absent in report	75
Follow-up plan absent in report	80

Of the 102 reports, 50 (49%) did not fully satisfy the referring physician; internists 49% and GP's 48%. Ten reports (10%) were considered to contain unnecessary information, and 43 reports (42%) were considered to contain insufficient information (3 reports (3%) fell in both categories). Most of the criticised reports contained only 1 inadequacy (32), the most criticised report contained 4. Reported inadequacies and their frequencies are listed in Table 3.

Table 3. Frequencies of inadequacies in the 102 studied endoscopy reports as reported by the referring internists, GPs and both.

Subject	Internists	GPs	All physicians
Indication absent, but desired	28%	21%	25%
Detail of findings insufficient	4%	3%	4%
Unclear whether findings can account for complaints	10%	6%	9%
Answer to clinical question insufficient	3%	0%	2%
Therapy plan absent, but desired	6%	12%	8%
Therapy plan present, but undesired	6%	12%	8%
Follow-up plan absent, but desired	12%	9%	11%
Follow-up plan present, but undesired	3%	3%	3%

Indication, findings and answer to clinical question

The indication for endoscopy was absent but desired in 26 reports (25%). As this item was absent in 46 reports (45%), more than half of the referring physicians stated that an indication should have been included. None of the respondents considered the reported findings to be too detailed. In 4 reports (4%) the referring physician considered the findings to be insufficiently described: in 2 cases the referring physician could not determine whether polyps had been removed completely, in one case the referring physician was not provided with a description of a process in the rectum, and in one case, a report on an endoscopy performed for anal bleeding did not mention the presence or absence of internal haemorrhoids. In 9 cases (9%), the physician was left with uncertainty as to whether the described findings could account for the complaints of the patient. These reports mostly described questionable abnormalities such as 'vague gastritis', 'reflux of bile' and 'some redness of the distal part of the esophagus'. Two reports (2%) were considered to have given insufficient answer to the clinical question. Unfortunately, both respondents failed to state which information should have been included.

Therapy and follow-up plan

In 8 reports (8%), the referring physician felt that a therapy plan should have been included. In 4 of these cases the referring physician did not know whether or, if so, which therapy should be given. In two reports, the referring physicians did not know whether therapy was indicated, but felt that the report correctly did not contain a therapy plan. In 8 reports (8%) an undesired therapy plan had been given. When a therapy plan was given (27 reports), none of the respondents disagreed with the suggested therapy.

A follow-up plan was absent, but desired by the referring physician in 11

reports (11%). In 6 of these cases the physician did not know whether or when follow-up had to take place. Undesired follow-up information was given in 3 reports (3%).

Discussion

In this study we assessed the opinions of physicians, who had referred patients to our endoscopy unit, on the content of the resulting endoscopy report. Similar surveys have been performed on radiology reports [1] and discharge letters [2]. Comparison of results, however, is difficult as terms such as 'poor or barely adequate', 'confusing' and 'not fully satisfactory' have different meanings. Secondly, some reported inadequacies may be specific for endoscopy, as, in the Netherlands, an endoscopy examination lies in the grey area between an additional examination such as radiology (describe what has been seen) and a true referral (describe also what has to be done). Therefore, inadequacies pertaining to a missing or undesired therapy or follow-up plan may be a reflection of a difference in opinion between endoscopist and referrer on the status of an endoscopy. Finally, in our study opinions of referring physicians were obtained when they had just received the endoscopy report, whereas in the other studies a general opinion on radiology reports was asked [1] or letters were judged by a panel [2]. Therefore, our study may have made deficiencies as experienced by the referrer more realistically apparent.

Despite the fact that results cannot be compared, our study shows that almost half of the endoscopy report were not to the full satisfaction of the referring physician. The most frequently mentioned shortcomings were: absence of indication, absence of follow-up, uncertainty whether findings could account for complaints and absence or undesired presence of a therapy plan. This in contrast to endoscopists who focused most remarks on insufficiently described findings [3].

What consequences do these results have for the contents of endoscopy reports?

As none of the respondents considered findings to be described in too much detail, there is no need beforehand to send referring physicians more concise reports than the reports used by endoscopists, even when reports become more detailed, e.g. through computerised reporting. As no differences could be found in the opinions between internists and GP's,

there is also no need to create different reports for these different groups of referring physicians.

For indications, we propose that they should always be included. Not only is this in co-accordance with the various guidelines for quality assessment [7], but there is also no argument not to include an indication.

A guideline with regard to information on whether findings can account for complaints of the patient, is more difficult to give. A referring physician obviously has a problem when he is uncertain whether or not further examinations are needed to clarify the patient's complaints. However, in most cases where the referring physician was left with such uncertainty, the report described rather unspecific findings, on which a consulted endoscopy expert could draw as little conclusion as the referring physician. So, in such cases, the best option will probably be to add a remark in the conclusion stating that it remains unclear whether the findings can account for the complaints.

As to inclusion of a therapy and follow-up plan, one could decide to always include such a plan whenever an abnormality is found. However, this would not satisfy physicians who indicated that they did not appreciate inclusion of such plans. Their response may be motivated by the fact that they may be held responsible whenever they do not comply with a plan and something goes wrong with the patient.

The ability of an endoscopist to fulfil the referrer's preferences depends on those preferences being apparent. If one holds the referrer responsible for the information that the endoscopist is to give, then, a therapy and follow-up plan are best always given, unless the referrer has made explicit that such information is not desired. However, one could also argue that it belongs to the endoscopists' responsibilities to facilitate the explicit formulation of the information preferences by the referring physician. In this light, a revised request/referral form, on which a referrer can clearly indicate which information he desires [10] would optimise communication between endoscopists and referring physicians.

Acknowledgement

We thank all physicians for their participation. This study was financially supported by Glaxo BV.

References

1. Clinger NJ, Hunter TB, and Hillman BJ. Radiology reporting: attitudes of referring physicians. Radiology 1988; 169: 825-6.
2. Westerman RF, Hull FM, Bezemer PD, et al. A study of communication between general practitioners and specialists. Br J Gen Pract 1990; 40: 445-9.
3. Moorman PW, Van Ginneken AM, Van der Lei J, et al. The contents of free text endoscopy reports: Inventory and peer review. Endoscopy 1994; 26: 531-8.
4. Stöltzing H, Birkner B, Lindbar R, et al. Computerunterstützte Dokumentation bei der oberen gastrointestinalen Endoskopie: Erfahrungen bei der Routineanwendung in drei Kliniken. Z Gastroenterol 1989; 27: 667-75.
5. Nguyen HN, Walker S, and Bode JCh. Routineeinzatz eines computer-unterstützten Dokumentationssystems für die Ösophago-Gastro-Duodenoskopie. Z Gastroenterol 1991; 29: 121-4.
6. Kuhn K, Gaus W, Wechsler JG, et al. Structured reporting of medical findings: evaluation of a system in gastroenterology. Methods Inf Med 1992; 31: 268-74.
7. Schapiro M. Computerization of endoscopic reports - an ASGE Proposal. Endoscopy 1992; 24 Suppl. 2: 478-80.
8. Gouveia-Oliveira A, Raposo VD, Salgado NC, et al. Longitudinal comparative study on the influence of computers on reporting of clinical data. Endoscopy 1991; 23: 334-7.
9. Aabaken L, Westerheim J, Hofstad B, et al. SADE database for endoscopic procedures: aspects of clinical use. Endoscopy 1991; 23: 269-71.
10. Scally PM. Design of a radiology request/referral form. Australas Radiol 1993; 37: 201-4.

Chapter 4

How often is large smaller than small?

Moorman PW [a,b]
Siersema PD [a]
de Ridder MAJ [c]
and van Ginneken AM [b]

[a] Department of Internal Medicine,
[b] Department of Medical Informatics,
[c] Institute Epidemiology and Biostatistics,
Erasmus University, Rotterdam, The Netherlands.

Published in shortened form as letter in
Lancet 1995; 345: 865

Abstract

In endoscopy reports, one third of the sizes of gastric ulcers are described using a non-numerical expression only. To study in which extent such expressions of size form a source of ambiguity, we sent a questionnaire to 222 Dutch physicians who perform endoscopies. They were asked to make the expressions 'small', 'medium', and 'large', when used to describe the size of a gastric ulcer, explicit on a numerical scale. Response rate was 71.2%. Overlap in the range of terms was large. In 31.1% 'large' did not exceed 'small'. As this may have clinical consequences, we recommend the use of numerical size estimates.

Introduction

Differences in the meaning attributed to non-numerical expressions of frequency and probability become explicit when physicians are asked to assign a numerical value to expressions such as 'usually' and 'likely' [1,2].

In a previous study [3], we found that in one third of the cases endoscopists describe the size of a gastric ulcer by a non-numerical expression only. In these non-numerical size descriptions we encountered as many as 11 different terms. To examine whether non-numerical size descriptions of gastric ulcers are ambiguous, and thus a potential for misunderstanding, we asked Dutch endoscopists to express descriptions of size in a numerical range.

Methods and Materials

We sent a questionnaire to the 222 Dutch gastroenterologists and internists, who routinely perform endoscopies. We asked these endoscopists to give the consecutive numeric range, that is the lower and upper limit, for the expressions 'small', 'medium', and 'large', when they use these expressions to describe the size of a gastric ulcer. These 3 expressions were chosen because a) they are the most often encountered terms to describe gastric ulcer size [3], and b) because the terminology proposed by the World Organization for Digestive Endoscopy (OMED) [4] suggests these terms as attributes of size, although they advocate that size should preferably be given in cm's.

To test whether the values given by the endoscopists depended on experience, or on the location where they had been trained in performing endoscopies, we also asked the endoscopists the number of years with endoscopy experience and the hospital in which they had their training.

For each term we determined the overall range (from minimum lower limit to maximum upper limit) and the 90% range: the range with the lower and upper extreme 5% eliminated. For the upper limit of 'small' (=lower limit of 'medium'), and the lower limit of 'large' (=upper limit of 'medium') we calculated the mean, standard deviation, range, median and mode. For these limits, we furthermore compared endoscopists who had 5 years or less experience with those who had more than 5 years of experience (t-test, and variance ratio test), and performed a variance analysis to assess the influence of training hospital.

In order to determine the overlap in 'small' and 'large', we compared the value of the upper limit of 'small' as given by each endoscopist to the value of the lower limit of 'large' of each of his 149 colleagues, and calculated the proportion of combinations in which 'large' did not exceed 'small'. When there is overlap in the ranges of small and large between two endoscopists, then there is disagreement in one direction only. Therefore, the total number of combinations between endoscopists is 150*149/2 (= 11175).

Results

In total 158 questionnaires (71.2%) were returned, of which 150 were valid. The 8 non-valid responses came from endoscopists who stated that they always described the size of gastric ulcers numerically, and thus were unable to fill in the questionnaire.

Figure 1 depicts the overall and 90% ranges of the terms 'small', 'medium', and 'large': e.g., the lowest value given for the lower limit of 'medium' was 0.2 cm, the highest upper limit given was 5 cm. An upper limit for 'large' was not given by 54 endoscopists.

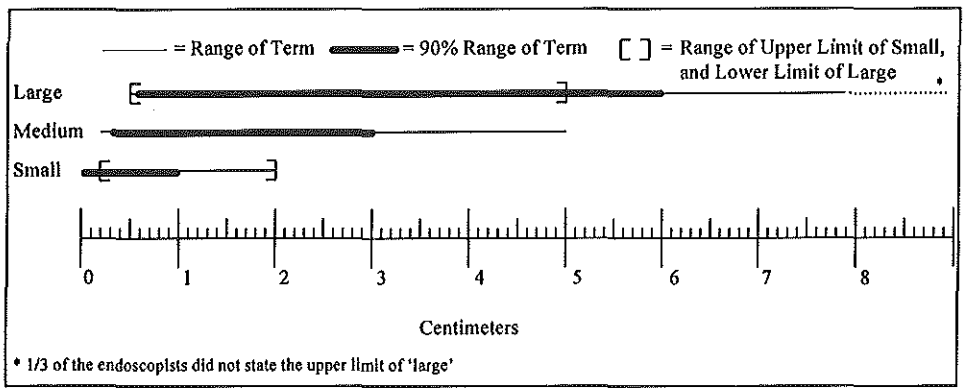


Figure 1. The ranges of the terms 'small', 'medium' and 'large' as they are used by endoscopists to describe the size of a gastric ulcer. The bold parts of the lines indicate the '90 % range', that is the range with the lower and upper extreme 5% eliminated. The range of the upper limit of 'small', and the lower limit of 'large' is given between the [ ]'s.

Table 1 lists the descriptives of the upper limit of 'small', and the lower limit of 'large'.

Table 1. Descriptives of the upper limit of small and lower limit of large.

	Mean	Std Dev	Min	Max	Median	Mode (nr of endoscopists)
Upper limit of small	0.75	0.33	0.20	2.00	0.75	1.00 (58)
Lower limit of large	1.71	0.74	0.50	5.00	1.70	2.00 (49)

There was no difference between the limits given by experienced (>5 years) and less experienced (<=5 years) endoscopists (t-test: upper limit of small: p=0.76, lower limit of large: p=0.35), nor in the variances between these 2 groups (variance ratio test: upper limit of small: p=0.43, lower limit of large: p=0.13). Furthermore, there were no differences between the different training hospitals (upper limit of small: p=0.835, lower limit of large: p=0.421).

In 3480 (31.1%) of the 11175 combinations, the lower limit of 'large' stated by endoscopist A did not exceed the upper limit of 'small' stated by endoscopist B.

Conclusions

When physicians read their own medical notes, they use their own reference to interpret terms, such as probably, usually, normal, and large. Problems may occur when the interpreter is not the same person as the author of these notes. In this study we made the possible ambiguity of non-numerical size descriptions explicit.

There was large overlap in the meaning that endoscopists assign to the 3 studied size expressions for gastric ulcers. Not mentioning a size at all seems to give as much (or little) information as using one of these expressions: the range of 'medium' fell completely within the ranges of 'small' and 'large'. Even when we eliminate the 5% most extreme values there remains overlap in 'small' and 'large'.

In 31.1% of the cases, an endoscopist may use the term 'small' to describe an ulcer, that is described as 'large' by his colleague. In such cases, non-numerical size descriptions may lead to sub-optimal patient care. It is possible that an ulcer is judged as large by one endoscopist, and on a consecutive examination as small by another in a situation where in reality the ulcer *increased* in size.

Although a numerical size *estimate* is never free of 'measuring' errors, they are less ambiguous than non-numerical expressions. Particularly,

because numerical size estimates seem to suffer more from a systemic underestimation than a random error [5,6].

Acknowledgement

This research was financially supported by Glaxo BV.

References

1. Bryant GD, and Norman GR. Expressions of probability: Words and numbers. N Engl J Med 1980; 302: 411.
2. Toogood JH. What do we mean by 'usually'? Lancet 1980 ; 17 : 1094.
3. Moorman PW, van Ginneken AM, van der Lei J, et al. The contents of free-text endoscopy reports: An inventory and evaluation by peers. Endoscopy 1994; 26: 531-8.
4. Maratka Z. *Terminology, definitions and diagnostic criteria in digestive endoscopy, 2nd ed. OMED database of digestive endoscopy.* Bad Homburg: Normed Verlag, 1989.
5. Fennerty MB, Davidson J, Emerson SS, et al. Are endoscopic measurements of colonic polyps reliable? Am J Gastroenterol 1993; 88: 496-500.
6. Vakil N, Smith W, Bourgeois K, et al. Endoscopic measurement of lesion size: improved accuracy with image processing. Gastrointest Endoscopy 1994; 40: 178-83.

Chapter 5

Descriptive features of gastric ulcers: Do endoscopists agree on what they see?

Moorman PW [a,b]
Siersema PD [b]
and van Ginneken AM [a]

[a] Department of Medical Informatics, Erasmus University,
[b] Department of Internal Medicine, University Hospital Dijkzigt,
Rotterdam, The Netherlands.

Gastrointest Endoscopy 1995; 42: (in press)

Abstract

Little is known about the interobserver variation between endoscopists on descriptive morphological features. This study describes the agreement among 10 endoscopists on their description of 12 morphologic features, using photographs of gastric ulcers, and on their eventual interpretation. The endoscopists used a form with predefined options for description. To analyze agreement, we assigned the descriptions given to one of three categories.

Kappa value was on average 0.36 for descriptive features, and 0.31 for interpretation. The proportion of endoscopists agreeing on descriptive features was on average 84%, and 81% on interpretations. The chance of an endoscopist describing all 12 morphological features of an ulcer on a photograph exactly the same as a colleague ranged from 4% to 46% (average 15%).

These results indicate a poor agreement between endoscopists in their translation of visual observations in descriptive terms. The positive correlation between agreement in description and interpretation (0.75, $p<0.05$) suggests disagreement in description as an important cause for disagreement in interpretation. We believe that making the meaning of descriptive terms explicit may improve agreement in description and in subsequent interpretation.

Introduction

Reporting plays an essential role in endoscopy. Endoscopy reports convey the findings of an endoscopy to the physician who requested the endoscopy, and serve as reference material for future examinations. The importance of reports is underlined by the development of guidelines for the contents [1,2] and the development of a standardized terminology [3].

Describing the same topics and using the same terminology, however, is no guarantee that endoscopists will describe identical findings in a similar way. Previous research has shown that endoscopic findings lack accuracy (e.g. size estimates suffer from an underestimation of up to 30% [4,5]), and that essential parts of endoscopy reports such as diagnoses and interpretations suffer from interobserver variation [6-8]. Interobserver variation is not unique to endoscopy. It is also reported in other clinical disciplines [9-10] and histology [11].

Most endoscopists will agree that an endoscopic diagnosis such as 'malignant gastric ulcer' is unreliable because of lacking accuracy - subsequent histology may prove otherwise [12] - and reproducibility - another endoscopist may classify the same lesion as benign. In contrast, little is known about the reliability of descriptive statements such as 'the ulcer has an irregular border'. Insight in such reliability is important as it provides information on:

- the probability that 2 endoscopists describe morphological descriptive features the same way,
- the a priori predictive value of a feature for the diagnosis of a lesion: e.g. when a sharply demarcated ulcer edge is present in 11 of 20 benign ulcers and in 2 of 20 malignant ulcers [13], accurate assessment of the predictive value of such statements is not possible when the reliability of a description of the ulcer demarcation is not known.

Determining the reliability of descriptive features in respect to the 'truth' is difficult, because defining the gold standard is virtually impossible; microscopy, for example, cannot state with more certainty whether the base of an ulcer is regular or irregular. Another measure of reliability, however, is the agreement between endoscopists on descriptive statements. In this study, we assessed the interobserver variation between endoscopists regarding descriptive morphological features and interpretation of what is considered a difficult endoscopic diagnosis: gastric ulcer.

Methods and Materials

Photographs, endoscopists and evaluation form

From our Gastroenterology Unit, we retrospectively obtained the ten most recent slides of gastric ulcers, which were of reasonable technical quality (i.e. sharpness, contrast). We made paper-prints (photographs) of these slides. Ten experienced endoscopists were asked to participate, and all agreed. Two endoscopists work in our university hospital, the remaining 8 practice in hospitals affiliated to the university.

We asked the endoscopists to evaluate the 10 photographs using a specially designed evaluation form. The form offered predefined options to describe 12 main morphological *features* of a gastric ulcer. To give an example: the feature 'shape' could be described by the options 'circular', 'oval', 'linear', 'serpiginous', and 'irregular'. The endoscopists were allowed to select more than one option per feature. For each photograph and for each feature, the endoscopist had the option to indicate that a reliable description of that particular feature was not possible. When an endoscopist indicated that bleeding stigmata were present, he could specify these by selecting one or more of the following: clot, visible vessel and / or active bleeding.

After describing an ulcer, the endoscopists were asked to give a diagnostic impression of that ulcer, using a 5 point scale ranging from 'possibly benign' to 'possibly malignant'. We will refer to this diagnostic impression as '*interpretation*'.

Data analysis

To analyze agreement, we grouped the descriptions given by the endoscopists into 3 categories (Table 1). The first category comprises descriptions that can be regarded as being contradictory to those in the second category. The third category constitutes the answers in which the endoscopist indicated that no reliable description for that feature could be given. When an endoscopist had described a feature using options in both category I and II, the description was assigned to the category II.

Agreement can be expressed in several ways. In our study we used the following: kappa, proportion of agreeing endoscopists, and the chance that individual endoscopists produce the same description of a photograph (chance of same description).

Table 1. The morphological features of gastric ulcers, the options for description on the evaluation form, and the categories to which they were assigned.

Subject	Options		
	Category I	Category II	III
Morphological features			
Shape	circular, oval, linear	serpiginous, irregular	npd*
Depth	superficial	medium deep, deep	npd
Base : regularity	regular	irregular	npd
Base : exudate	absent	present	npd
Border : elevation	flat	partially raised, / completely raised	npd
Border : regularity	regular	irregular	npd
Border : undermining	absent	present	npd
Surrounding mucosa : color	normal	red, pale	npd
Surrounding mucosa : swelling	absent	present	npd
Surrounding mucosa : nodules	absent	present	npd
Demarcation from surroundings	sharp	vague	npd
Stigmata of bleeding	absent	present	npd
Diagnostic impression			
Interpretation (From possibly benign to possibly malignant)	1, 2	4, 5	3

* npd = not possible to give a reliable description

Kappa. For agreement between two endoscopists on the three categories of an ulcer feature, Cohen's kappa [14-16] can be calculated by the formula

$$\kappa = \frac{Po - Pe}{1 - Pe}$$

where Po is the observed agreement and Pe is the agreement expected by chance. (The calculation of Po and Pe is described in the appendix.) The overall kappa value for all endoscopists, may be calculated by averaging all pairwise calculated kappa's. When Pe equals one, kappa cannot be calculated. Kappa can range from -1 to 1, and is constructed to be zero when the obtained agreement can be entirely attributed to chance. The interpretation of kappa values is somewhat subjective, but kappa values above 0.75 are considered to represent excellent agreement, and values below 0.40 poor agreement [17]. Although kappa is a generally accepted measure, a difficulty in the interpretation is that kappa is also affected by the presence of bias between observers (e.g. when endoscopists assign observations predominantly to one category) and by the distribution of data across the categories [18].

Proportion of agreeing endoscopists (PAE). In most studies assessing interobserver agreement, only two or three observers are involved. The ten observers in our study permit to express agreement also in the proportion of agreeing endoscopists (PAE). As PAE signifies the chance that an endoscopist would describe a feature the same as a colleague, it is a more intuitive and illustrative measure of agreement than kappa. When the answers are dichotomized in contradictory statements, then the proportion of endoscopists agreeing on that topic is defined as

$$\mathrm{PAE} = \frac{100x}{(x+y)}\%$$

where x endoscopists (the largest proportion) state option X, and y endoscopists state Y. PAE ranges from 50% (half of the endoscopists state X, the other half state Y) to 100% (all endoscopists state X or Y). When, for example, PAE is 80% then every fifth endoscopist has stated the contrary of the other four. Mean PAE for a feature was calculated by averaging the PAE of that feature on every photograph.

In this study, differences in the values of kappa and PAE originate from the fact that 1) unlike PAE, the attribution of chance agreement is eliminated in the calculation of kappa, and 2) kappa also includes disagreement among endoscopists on whether or not a reliable assessment could be given.

Chance of Same Description (CSD). When we assume that PAE's for the features of a given photograph are independent, then we can calculate the chance that, given a description, a second endoscopist would give exactly the same description for that photograph by the formula

$$\mathrm{CSD} = \prod_{i=1}^{N} \mathrm{PAE}i$$

where N is the number of features, in our case 12. Note that the assumption of independence does not relate to independence of the appearance of features, but to *independence of agreement.* We think that such an assumption is acceptable, although we acknowledge the fact that some degree of correlation between the agreement on various features may exist.

In addition, we calculated *Spearman rank correlation coefficient* to test whether agreement on interpretation ($PAE_{interpretation\ of\ photo\ x}$) correlates with agreement on description ($CSD_{photo\ x}$).

Results

Descriptive features. Kappa values for descriptive features ranged from 0.06 to 0.59, and was on average 0.36 (Table 2). Highest kappa value was found for the feature describing whether or not the ulcer was superficial. Poor agreement among the endoscopists was found for the description of the 'shape' (regular vs. irregular), 'elevation of the border', and all features concerning the surrounding mucosa.

Table 2. Kappa value and proportion of agreeing endoscopists on descriptive gastric ulcer features and interpretation.

Feature	Kappa	PAE (range)
Shape	0.38	86% (60-100)
Depth	0.59	88% (70-100)
Base : regularity	0.40	82% (60-100)
Base : exudate	-*	99% (90-100)
Border : elevation	0.33	78% (60-100)
Border : regularity	0.41	81% (55-100)
Border : undermining	0.46	91% (60-100)
Surrounding mucosa : color	0.23	79% (50-100)
Surrounding mucosa : swelling	0.06	69% (50-87)
Surrounding mucosa : nodules	0.20	81% (50-100)
Demarcation from surroundings	0.44	87% (60-100)
Stigmata of bleeding	0.43	88% (57-100)
Average	0.36	84%
Interpretation	0.31	81% (55-100)

* Kappa could not be calculated (Pe = 1)

The proportion of agreeing endoscopists (PAE) on gastric ulcer features ranged from 69% to 99%, and was on average 84% (Table 2). For the features 'presence of exudate' and 'undermining of the border', the PAE was larger than 90%. PAE was less than 80% for the features 'flat or elevated' border, 'normal or abnormal color' of surrounding mucosa, and 'presence of swelling' of surrounding mucosa. For the absence or presence of 'active bleeding', 'a clot' and 'a visible vessel' PAE's respectively were 95%, 92% and 79%.

Interpretations. The average proportion of endoscopists agreeing on the interpretation of a photograph of a gastric ulcer was 81%. Kappa value for the interpretation was also slightly below the average of descriptive features, namely 0.31.

Chance of same description. The chance of an endoscopist describing all 12 morphological features of an ulcer on a photograph exactly the same as a colleague (CSD) ranged from 4% to 46% (average 15%).

Correlation between description and interpretation. Spearman correlation between CSD of a photograph and the average PAE on the interpretation was 0.75 ($p < 0.05$).

Discussion

This study describes the agreement among 10 endoscopists on their description of 12 morphologic features using photographs of gastric ulcers, and on their eventual interpretation.

The average kappa value for interpretation of gastric ulcers (0.31) indicates poor agreement between endoscopists, and reflects that interpretation is a complex process. Other studies [6-8] have already shown an average to low agreement on endoscopic diagnoses, but do not permit any conclusion on the nature of the low agreement. Two hypotheses may account for the lack of agreement on interpretations.

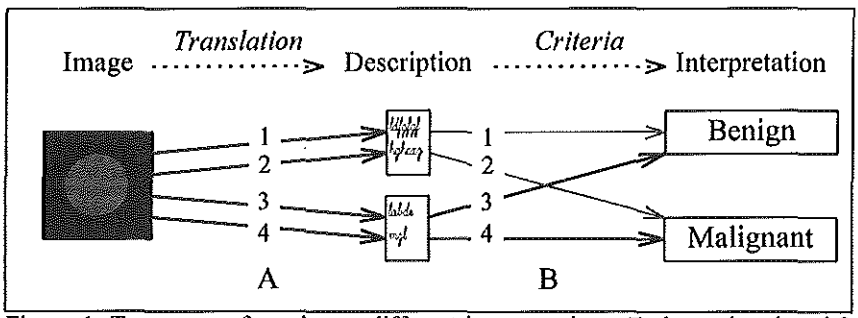


Figure 1. Two ways of coming to different interpretations. Endoscopists 1 and 2 give the same description of the lesion, but this description differs from the one given by endoscopists 3 and 4 (A). Starting from the same description, endoscopist 1 and 2 apply different criteria, and therefore arrive at a different interpretation (B).

First, endoscopists may differ in their criteria about what constitutes a

malignant or benign ulcer (Figure 1, part B). Second, they agree in their criteria, but fail to translate their visual observation in equal descriptive terms (Figure 1 part A). If this second hypothesis dominates, endoscopists produce different descriptions on the basis of a given image of a lesion, but would arrive at the same interpretation on the basis of a given description of a lesion.The correlation we found between agreement on description and agreement on interpretation (0.75, $p < 0.05$) supports the second hypothesis. This correlation indicates that where endoscopists give the same description, they also tend to arrive at the same interpretation, while giving different descriptions, their interpretations also differ. Disagreement in description thus accounts for low agreement on interpretation.

Disagreement in description may play an important role in daily practice: it is comparably low to agreement on interpretation (0.36 vs. 0.31). The proportion of endoscopists agreeing on a descriptive feature was on average 84%, which signifies that if 20 endoscopists were to assess a single feature, then 17 would state e.g. regular and 3 irregular, or vice versa. On the assumption of independence of agreement on features, the chance that two endoscopists describe all 12 features in the same way (CSD) becomes very small, 15%.

These are important observations, as it also calls for caution in the interpretation of relations between descriptive endoscopic features and other observations, e.g. histological diagnoses. It is not inconceivable that the poor correlation between, for example, an irregular base and a (histologic) malignant ulcer is largely due to endoscopists failing to agree what constitutes an irregular base. In general, clinical studies relying on descriptive morphological features in endoscopy (such as comparing effects of medications) pay little or no attention to the potential interobserver variation at the observational level.

How well do the discussed measures of agreement reflect reality?

As compared to clinical practice, factors that may have caused under- and overestimation of kappa, PAE and CSD need to be considered.

Underestimation of agreement. As photographing gastric ulcers is not a standard procedure, photographs of 'interesting' or 'difficult' ulcers may be over-represented in this study. Furthermore, the endoscopists made their descriptions from 2-dimensional photographs and did not actually perform endoscopies themselves, depriving them from looking at the ulcers from different angles (pseudo 3-dimensional view). However, the endoscopists

had the option to indicate that no reliable description of a feature could be given; an option that was not used very often. In addition, our data do not indicate that this limitation played an important role, as agreement on '2 dimensional features' did not differ much from agreement on '3-dimensional features'. In fact, it was surprising that the highest kappa value was obtained for the assessment of depth.

Overestimation of agreement. Agreement in real practice may even be lower than the agreement we found, as in our study the endoscopists were confronted with the fact of a present ulcer, and kappa for the identification of the presence of an ulcer has been reported to be only 0.7 [7]. Furthermore, as we categorized the descriptions given, endoscopists may also disagree within the same category. For example, the category 'abnormal color of the surrounding mucosa', included the options 'red' and 'pale'. Disagreement within this category 'abnormal color' was still 10%.

What could be done to improve agreement between endoscopists on descriptive features?

One could start by making the meaning of terms explicit. Although a statement such as 'irregular border' may seem unambiguous, we found it could have two meanings namely: the elevation is irregular in height, or the elevation is irregular in width. Thus, the same assessment is given in differing situations. Making the meaning of descriptive terms explicit thus is important, but does not necessarily require a pure linguistic approach. When endoscopists describe features, it is likely that they use conceptual reference images; they compare what they see with an image in their memory. Providing the endoscopists with equal descriptions of reference images may therefore already improve agreement in description. Although this option may be realizable in an educational setting, it seems unpractical in the daily clinical setting, unless we make use of computer tools. In the future, we may envision an endoscopic reporting program where, upon selecting a term, the user is provided with images in which that term is visualized.

We believe that reducing disagreement in endoscopic descriptions will increase the value of endoscopy reports in practice and for research. Meanwhile, it seems important that clinical studies should strive to formulate descriptive features as explicit as possible, and should state the number of endoscopists that have participated in the study. For clinical

practice, we believe that adding a photograph of the observed lesion to the report, will help to ensure that the correct message is conveyed to the referring physician, and to the endoscopist performing the follow-up examinations.

Acknowledgment

We thank Dr. M. van Blankenstein and Dr. J. van der Lei for their help in the preparation of this document. This research was financially supported by Glaxo BV.

Appendix

Suppose two raters classify N subjects as belonging to one of three categories. The result can be arranged in a 3 x 3 table as follows:

		observer 1			
		I	II	III	Total
observer	I	a	b	c	j_1
	II	d	e	f	j_2
	III	g	h	i	j_3
	Total	k_1	k_2	k_3	N

where, e.g., d is the number of subjects assigned to category I by observer 1, and to category II by observer 2. The proportion of observed agreement is Po = (a+e+i)/N. Agreement expected by chance is Pe = $(j_1k_1 + j_2k_2 + j_3k_3)/N^2$.

References

1. American Society for Gastrointestinal Endoscopy. Quality assurance of gastrointestinal endoscopy: an information resource manual, 1989.
2. Schapiro M. Computerization of endoscopic reports - an ASGE Proposal. Endoscopy 1992; 31 Suppl. 2: 478-80.
3. Maratka A. *Terminology, definitions and diagnostic criteria in digestive endoscopy. 2nd ed. Omed database of digestive endoscopy.* Bad Homburg: Normed Verlag, 1989.
4. Fennerty MB, Davidson J, Emerson SS, et al. Are endoscopic

measurements of colonic polyps reliable? Am J gastroenterol 1993; 88: 496-500.
5. Margulies C, Krevsky B, and Catalano MF. How accurate are endoscopic estimates of size? Gastrointest Endosc 1994; 40: 174-7.
6. Bendtsen F, Skovgaard LT, Sørensen TIA, and Matzen P. Agreement among multiple observers on endoscopic diagnosis of esophageal varices before bleeding. Hepatology 1990; 11: 341-6.
7. Bytzer P, Havelund T, and Møller Hansen J. Interobserver variation in the endoscopic diagnosis of reflux esophagitis. Scand J Gastroenterol 1993; 28: 119-25.
8. Woolf GM, Riddell RH, Irvine EJ, and Hunt RH. A study to examine agreement between endoscopy and histology for the diagnosis of columnar lined (Barrett's) esophagus. Gastrointest Endosc 1989; 35: 541-4.
9. Koran LM. The reliability of clinical methods, data and judgments. N Engl J Med 1975; 293: 642-6 and 695-701.
10. Komaroff AL. The variability and inaccuracy of medical data. Proc IEEE 1979; 28: 1196-207.
11. Christensen AH, Gjørup T, Hilden J, et al. Observer homogeneity in the histologic diagnosis of Helicobacter pylori. Scand J Gastroenterol 1992; 27: 933-9.
12. Dekker W, and Tytgat GN. Diagnostic accuracy of fiberendoscopy in the detection of upper intestinal malignancy. A follow-up analysis. Gastroenterol 1977; 73: 710-4.
13. Gabrielsson N. Benign and malignant gastric ulcers. Evaluation of the different diagnostics in roentgen examination and endoscopy. Endoscopy 1972; 4: 73-83.
14. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Measurem 1960; 20: 37-46.
15. Fleiss JL. *Statistical methods for rates and proportions*. New York: J Wiley & Sons, 1981.
16. Schouten HJA. Statistical measurements of interobserver agreement (Thesis). Rotterdam: Erasmus University, 1985.
17. Landis JR, and Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977; 33: 671-9.
18. Byrt T, Bishop J, and Carlin JB. Bias, prevalence and kappa. J Clin Epidemiol 1993; 46: 423-9.

Chapter 6

A model for structured data entry based on explicit descriptional knowledge

Moorman PW [a,b]
van Ginneken AM [a]
van der Lei J [a]
and van Bemmel JH [a]

[a] Department of Medical Informatics, Erasmus University,
[b] Department of Internal Medicine, University Hospital Dijkzigt,
Rotterdam, The Netherlands

Meth Inform Med 1994; 33: 454-63

Abstract

Clinical narratives in patient records are usually recorded in free text, limiting the use of this information for research, quality assessment, and decision support. This study focuses on the capture of clinical narratives in a structured format by supporting physicians with structured data entry (SDE). We analyzed and made explicit which requirements SDE should meet to be acceptable for the physician on the one hand, and generate unambiguous patient data on the other. Starting from these requirements, we found that in order to support SDE, the knowledge on which it is based needs to be made explicit: we refer to this knowledge as descriptional knowledge. We articulate the nature of this knowledge, and propose a model in which it can be formally represented.

The model allows the construction of specific knowledgebases, each representing the knowledge needed to support SDE within a circumscribed domain. Data entry is made possible through a general entry program, of which the behavior is determined by a combination of user input and the content of the applicable domain. We clarify how descriptional knowledge is represented, modeled, and used by the data entry to achieve SDE, which meets the proposed requirements.

1. Introduction

The increased complexity and volume of information per patient has made the shortcomings of traditional paper-based records apparent [1]. These shortcomings include: incomplete, illegible, and inaccurate information, lack of standardized terminology, poor organization of the information, and limited availability, leading to suboptimal use of the information in the medical record.

The development of computer-based patient records is increasingly gaining attention. Early efforts focused predominantly on administrative, financial, and management requirements. Recording clinical data in a structured, coded fashion was typically confined to laboratory data, medication, and diagnoses (GEMISCH [2], TMR [2-3], COSTAR [4], Regenstrief [5], ELIAS [6]). The areas that are most difficult to capture in coded form - those parts of the medical record which involve descriptive information (clinical narratives) - were left to free text.

In this study, we explore the feasibility of minimizing free-text narratives in favor of clinically relevant structured data, suitable for multiple purposes: patient care, quality assessment, decision support and critiquing, and research. We use the word '*findings*' to denote any descriptive information, obtained during a patient encounter, regarding the history of present complaints, systemic inquiry and physical examination, as well as findings of additional examinations such as endoscopy.

In the following, we analyze requirements for the acquisition of structured data, and argue that a specific type of knowledge has to be made explicit to meet those requirements. Furthermore, we introduce and discuss a model in which such knowledge can be explicitly represented and used to support structured data entry.

2. Basic Considerations and Scope

To obtain structured medical data, two main methods may be distinguished. The first method is extraction of structured data from free text with natural-language-processing (NLP) techniques, the second is direct entry of data in a structured format, structured data entry (SDE). Extracting structured data from free text has the advantage that it is compatible with the current reporting routine; the method can thus be used on any free-text data in existing electronic records. Ideally, NLP can make

as much content explicit as a human reader, other than the author. In contrast, SDE performed by physicians themselves [7,8] offers the possibility to improve the quality of the data [9,10]. Therefore, we consider SDE as the preferred method for obtaining structured data.

To be acceptable for physicians, we believe that structured data entry should, at least, meet the following requirements:

- SDE has to provide the physician sufficient *expressive power* to describe in detail what he considers clinically relevant information.
- SDE has to correspond intuitively to the physician's *usual method of working*; applications that do not significantly change routine patient care are more likely to be accepted [11].
- SDE has to be *flexible*. The physician should be given freedom to determine the order and the degree of detail of what he describes.
- The entered data should be presented in a *predictable* order so that, when browsing through the data, a physician knows where to expect specific information.
- The time needed for structured entry of data should *not, or barely, exceed the currently needed reporting time.*

The data obtained by SDE also have to meet certain requirements:

- The program by which physicians enter the data should stimulate *completeness*: the collection of as many structured data about as many findings as possible.
- The patient data should be *unambiguous*. We call terms ambiguous when more information than the term alone is required to enable a unique medical interpretation. We call the information that makes a term unambiguous *context*. Especially when using data for various purposes, context needs to be ensured.

3. Identifying Descriptional Knowledge

Fast and flexible data entry, required to make SDE acceptable for physicians, is in conflict with obtaining extensive and unambiguous data. Furthermore, physicians will conform to a pre-defined structure and terminology only if the benefits outweigh the efforts imposed by SDE [12]. As the benefits will not always immediately be clear to the physician, and completeness and unambiguity are essential prerequisites for the use of clinical data for purposes such as research, we need to focus on the

practical feasibility of acceptable SDE. Acceptability will be enhanced if physicians do not lose valuable time by entering unnecessary information, and when trivial matters can be deduced by the system itself. Such 'intelligent' structured data entry can only be achieved when it is based on knowledge about the scope and meaning of the information that is to be captured. However, *what kind of* knowledge is needed? In the following section we make this knowledge explicit by describing the inferences that have to be made in order to support flexible SDE. Thereafter, we examine to what extent such knowledge, which we call *descriptional knowledge*, is present in existing knowledge representations and existing SDE applications.

3.1 What should Descriptional Knowledge make Explicit?

Terms that physicians are likely to use. When a physician documents his findings in a paper medical record, he uses knowledge. One physician may describe a cardiac murmur by its loudness, location of maximal loudness, and the phase of the heart-cycle in which it occurred. Another physician may not immediately remember to describe the murmur's loudness, but may add radiation to the description. By providing physicians a list of all terms that describe a given term, it may remind them of what they have forgotten, thus stimulating them to enter more data. Descriptional knowledge should *enable the identification of terms that physicians are likely to use* in the description of a given term. We will call the terms that a physician is likely to use in the description of a given term A, the *descriptors* of A.

Descriptors of terms in their context. Descriptors of a term do not depend on that term alone. By which terms a given term is likely to be described also depends on the circumstances in which it is used. When using 'penicillin' in the description of a therapy in the patient record, it is likely, and desirable, that thereafter the dosage is described. When 'penicillin' is used to describe an allergy of the patient, then a description of the dosage is irrelevant. We refer to these circumstances as *context*. Thus, the context in which a term is described also determines its descriptors. Therefore, descriptional knowledge should *enable the identification of the descriptors of a term in a given context.*

Contextual ambiguity. Context is not only essential for determining the descriptors of a term, but also for their unambiguous recording. A single note 'irregular pulse' raises the question whether it was reported by the patient or observed by the physician. Reliable interpretation is only possible when a statement is stored in its explicit context. Descriptional knowledge should enable to infer *whether a term may be used in one or more contexts.* The presence or absence of ambiguity can then be derived from the possible contexts, and the amount of context that is provided. The physician can be prompted to add the disambiguating information. At the same time, we increase the potential for flexible data entry: a physician does not have to know in advance how to access the term he wishes to describe.

The meaning of context for descriptions. Context also plays an important role in the ability to detect inconsistencies. Seemingly, 'an irregular pulse' cannot co-occur with 'no irregular pulse'. This is true within one context, e.g., patient history. However, if one of these statements applies to the patient history and the other to the physical examination, there is no contradiction. Hence, we need to be able to determine whether or not a statement, added to the description of a term, is described within the same context. For expressive power, however, it is required that some terms can be used more than once in a given context, each with its own description: three different ulcers in the stomach may require their own description.

Thus, descriptional knowledge should enable contextually dependent structured data entry, by which formally represented, unambiguous patient data may be obtained. Descriptional knowledge describes *where and how* entities can possibly be used. For instance, for a blister such knowledge would comprise:

- descriptors are size, color, and location,
- possible contexts are patient history and physical examination,
- for a given blister it must be specified which context applies,
- in each context multiple descriptions may occur,
- within one context, absent and present may not co-occur.

3.2 Other Studies Related to Structured Data Entry

In the preceding section, we have identified the tasks that descriptional knowledge should support, and have shown that physicians make use of this knowledge for writing findings in paper-based records and for

interpreting findings when reading notes. In this section, we discuss to what extent descriptional knowledge is present in several existing systems.

Medical vocabularies, such as UMLS [13] and SNOMED [14], are useful for knowing which terms to use; however, they do not specify what the descriptors of terms are: e.g., of an ulcer, its base and border may be described.

The work described in the PEN and PAD and the GALEN projects [15-20] focused on knowledge needed to infer whether a statement is 'sensible'. However, knowledge that defines what is sensible to say, does not define what is likely to be said. When 'diabetes' is mentioned in a medical record it is unlikely that a physician would want to state that it is an endocrine disease. Although relations defined in the Galen Representation and Integration Language (GRAIL), and its predecessor Structured Meta-Knowledge may be used to support SDE, we believe that more is needed for flexibility, e.g., to reduce the number of options.

When looking at the SDE applications described in literature [20-23], it is remarkable how little explicit attention is given to how the underlying knowledge model can functionally support those aspects that are important for SDE. We will discuss some implicit assumptions and explain what the consequences may be.

A similarity between many SDE applications [20-23] is that the physician needs to describe findings from the perspective of a problem. After selecting a list of problems, the physician may describe the findings related to these problems. The rationale behind this approach is that it limits the choices for the physician to relevant findings only, instead of offering the physician an overwhelming number of choices. However, the decision for such an approach has several consequences. In daily practice, findings not directly related to a problem may also be relevant: documenting the absence of a cardiac murmur is also important in non-cardiological patients. Yet, it is unclear in some applications how findings that are not related to a problem of a patient can be recorded.

It is also often not made explicit how a given finding described from the perspective of different problems is handled. Possibly, the program assumes that a finding is unique, irrespective of the problem-context: shortness of breath described from the problem 'cough' has to be the same as shortness of breath described via 'chest pain'. The assumption that a finding can only occur once in the description of a patient encounter has far-reaching consequences for flexibility and power of expression. We will illustrate this

with an example from Ivory [22-23]. In Ivory, 'cervical lymph nodes - left' and 'cervical lymph nodes - right' are both findings; both may be described by the modifier 'enlarged', which can take the values 'not specified', 'yes' or 'no'. Left and right side lymph nodes need to be described separately. This may be necessary, when a patient has the problem 'fever', but in daily practice it would be more convenient to describe both sides at once, and only mention laterality in case of something abnormal on one or both sides. Moreover, as the finding 'cervical lymph nodes - left' may occur only once, the physician cannot separately describe different enlarged lymph nodes on one side, but is restricted to describe the most prominent one. It remains unclear whether these are carefully chosen characteristics, or shortcomings of the underlying model.

The approach chosen in Ivory - findings described by a set of modifiers - is characterized by restricted depth and, therefore, does not allow for theoretical domain completeness [24]. It results in a limited descriptive level of detail, which has practical consequences: of an ulcer found during endoscopy, an endoscopist may want to describe, amongst others, its base and border, and of the base and border, the color and regularity [25]. Attempting to represent such multi-layered descriptions in the finding-modifier approach, may result in an unpractically large number of possible selections.

Some assumptions may also have consequences for unambiguity of obtained data. In a reporting system for ultrasound [26], for example, diagnostic criteria are used in descriptive options. A physician may, after he has stated the presence of a cyst in the ovary, describe its size by selecting between a cyst larger or smaller than 12 mm. One has to realize that by doing so, the size is recorded in relation to the diagnostic criteria valid at the time of recording. When these change it will be necessary, though probably impossible, to adapt all data to new insights, when the actual size is not available.

In summary, although several authors have addressed the issue of SDE of findings, they fail to describe a general strategy. We believe that the explicit identification of the required knowledge for SDE is an important prerequisite for understanding of and obtaining insight into the problems related to flexible and expressive SDE. We are not suggesting that the applications and knowledge representation schemes described could never enable data entry to conform with the requirements outlined in section 2,

but only that at present they do not contain explicit models of descriptional knowledge.

4. Representation of Descriptional Knowledge: General Approach

We propose an abstract model that provides a specific notation to represent descriptional knowledge - we will refer to this model as the *meta-model*. This meta-model provides the terms and relationships for expressing descriptional knowledge for a given application area; using the terms and relationships provided by the meta-model, we are able to create specific *descriptional knowledgebases* that contain the required descriptional knowledge for a given medical domain. The goal in developing our meta-model for descriptional knowledge was to support physicians in performing structured data entry; that is, the ability to employ such knowledge during data entry. Based on the terms and relationships of the meta-model, we developed a *general data entry program* that, given a descriptional knowledgebase, allows semantically correct structured data entry (Fig. 1). The result of that structured data entry are *patient data.*

In the following sections we first describe the meta-model. We subsequently describe how the terms and relationships provided by the meta-model are used to construct knowledgebases for specific domains. Finally, the general data-entry program and the patient data resulting from using that data-entry program are discussed.

4.1 The Meta-Model: Concepts, Relations, and Properties

The basic entity in medical texts is the *'concept'*. According to the CEN [27] and ISO [28,29], *concept* is a unit of thought, whereas *term* is a unit of language. We will use the notation [ ] to indicate a 'concept'; [X] should be read as 'the concept X'.

In a descriptional sense, concepts have in common *that they may be described further by a finite number of other concepts*. To represent that a concept may be the descriptor of another concept, we define relations between concepts; we use the notation < > to denote a relation. As a relation, in principle, means 'may be described by', the relation is directed from 'parent' (the concept to be described) to 'child' (the descriptor). In our model, we distinguish six different types of relations: <Has_Feature>, <Has_Specialization>, <Preset_Choice>, <Exclude_Choice>,

<Refers_To>, and <Has_Value>. These relations will be described in Section 5.1.

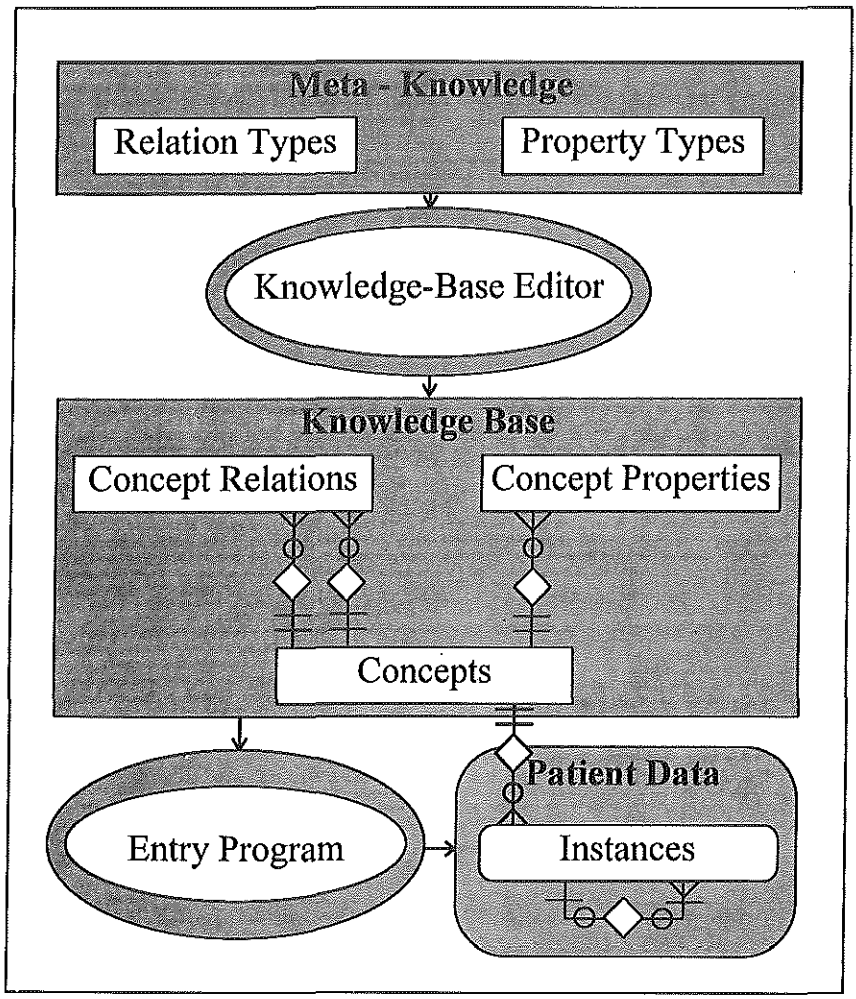


Figure 1. The model and its components. On the meta-level, the types of relations and properties that concepts may have are defined. Using the knowledgebase editor, descriptional knowledgebases may be created. In such a knowledgebase are defined which concepts may be used in a domain, which relations with other concepts, and which properties a concept has. On the basis of a knowledgebase, the entry program produces the patient data. Simplified entity-relationship diagrams are shown for the various components.

A concept can be the descriptor of more than one other concept; for example, [regularity] may be used in the description of [base] and [border]

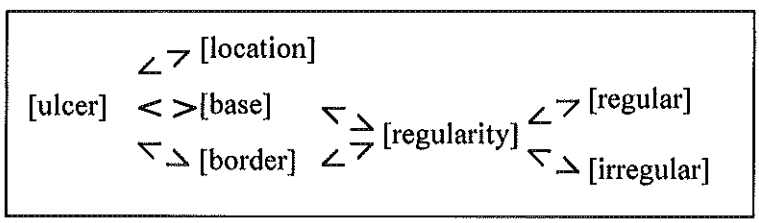
of an ulcer. Concepts and relations thus form a semantic network. Since [regularity] may as well be used to describe [base] and [border] of an ulcer, we cannot know which [regularity] is meant, when we do not know under which concept it was described. In our model, *the paths, in the opposite direction of the relations, represent the possible contexts in which a concept may be used in a medical record.* In addition, *a concept will only be unambiguous when it is explicit which of the possible paths applies.* The only exception is the *topnode* of the network; it has no parents. The topnode is the upper limit of the necessary context to make all other concepts unambiguous. It denotes the domain of all concepts in the network.

The relations between concepts and their descriptors determine how concepts may be described in increasing detail. However, concepts may also have certain intrinsic properties which determine how a concept may be used in a medical record. For example, [ulcer] differs from [mucosa] in the sense that there may be several [ulcer]s in the stomach, but only one [mucosa]. To represent such intrinsic properties of concepts, we distinguish four different properties: 'multiple', 'combination', 'absent', and 'normal'. These will be discussed in section 5.2. The notation {} will be used when referring to such a property.

4.2 Building a Descriptional Knowledge-Base

Given the relations and properties defined in the meta-model, the construction of a descriptional knowledgebase (KB) for a specific medical domain involves creating a thesaurus of concepts, defining relations, and assigning the correct properties to concepts.

To illustrate a descriptional KB network, or parts of it, we use a conceptual graphs notation [30]. Our graphs essentially represent which descriptors concepts have, and in which context concepts may be used. To refer to the specific use of conceptual graphs for the representation of descriptional knowledge we call these *'contextual graphs'*, such as:



We want to emphasize that we model *descriptional knowledge only.* That

is, we model that a physician may record that a patient has pale feces and colic-like abdominal pain. In terms of our model, colic-like abdominal pain is not linked to feces color, although they may coincide in case of gallstones.

4.3 Use of the Descriptional KB: Data Entry

The data-entry program enables the physician to describe the findings to be included in the patient record. The physician describes findings by selecting a series of concepts.

Starting with the topnode, the data-entry program reads from the descriptional KB which descriptors the concept has, and presents these for selection. The program infers which descriptors apply in a given context based on the already selected concepts and the types of relations between these concepts. The physician selects from the presented alternatives, and the cycle can be repeated for the selection; this selection has now become the concept to be described. We refer to the concept to be described as *current concept.* The sequence of selected concepts, from topnode to current concept, is called *concept path.* The data entry program allows the physician to go back along the concept path, and to select another alternative for description (see also Fig. 2).

The physician is not constrained to start at the topnode of the descriptional KB. At any point, the physician is able to select concepts 'in the middle of the KB', by entering (parts of) concept names. Because all paths by which a concept can be reached are represented in the KB, the entry program is able to determine all possible contexts of a concept. In case of more than one possible context for the selected concept, the program will first try to deduce the proper context. When a physician, for example, indicates that he is going to describe [pulse] after having made explicit that he is describing the [physical examination] (present in the concept path), then it is clear which context applies to [pulse]. When the physician tries to describe [pulse] without having previously indicated the context, the program will ask the physician to provide this context: the program prompts the physician to indicate whether he means to describe the [pulse] found during the [physical examination] or the [history].

In section 5 we describe in more detail the meaning of the different relations and properties for the data entry program.

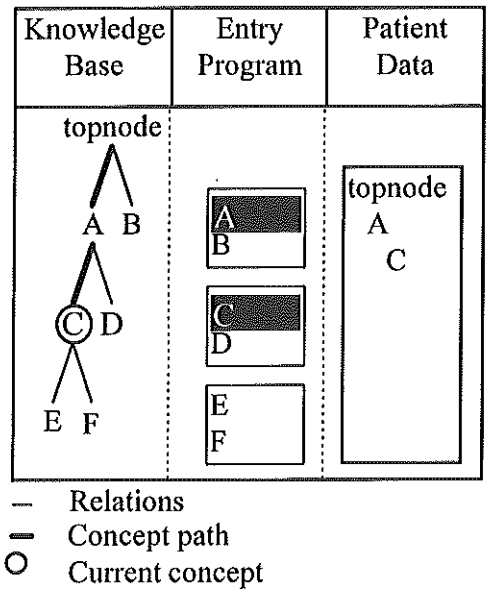


Figure 2. The knowledgebase is used by the entry program to present - for any selected concept - a menu with relevant options to be described next. In this example A and C have been selected, and thus are in the concept path. The menu with 'E F' presents the descriptors of the current concept (C). Selected concepts (patient data) have been instantiated, and represent the findings.

4.4 Patient Data

When describing a finding, the physician selects concepts sequentially; each selection adds to the context of the next concept. The *whole path of selected concepts represents a finding*, for example:

```
[topnode findings] <>
    [physical examination] <>
        [examination heart] <>
            [auscultation heart] <>
                [murmur] <>
                    [loudness] <>
                        [grade II]
```

Whereas the KB defines what possibly can be described, the patient data represent what actually applies to the patient. Since a concept may be used

in the description of several findings, the concepts are *instantiated in their context.* An instance, therefore, can only have one parent; instances in the patient data are thus represented as a tree structure in which the context of an instance is always unique. Figure 3 shows an example of the difference between concepts and instances.

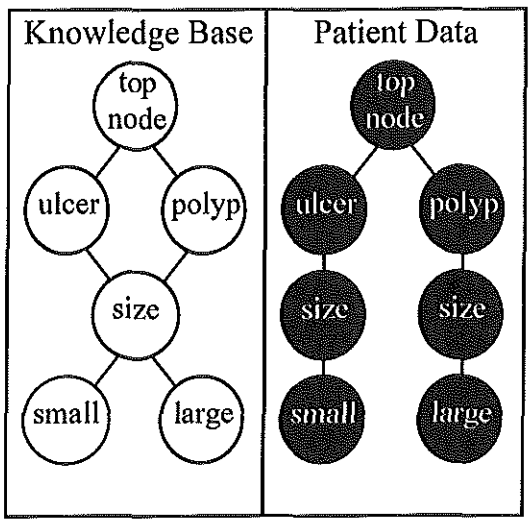


Figure 3. Example of the difference between concepts in the knowledgebase and instances in the patient data.

As relations between concepts in the KB serve only as information to support intelligent data entry, it is not required to record the original KB relation type between instances in the patient data: *relations between instances solely represent context.*

5. The Meta-Model's Components and their Meaning in Data Entry

In the preceding sections we have identified general characteristics of descriptional knowledge, and described in general terms how descriptional knowledge can be used for SDE while ensuring semantically correct patient data. In this section we describe in detail how each relation and property in our meta-model is used in constructing a descriptional KB, and how their semantic meaning is used by the entry program.

5.1 Relations

We will first discuss relations: <Has_Specialization>, <Has_Feature>, <Preset_Choice>, <Exclude_Choice>, <Refers_To>, and <Has_Value>.

5.1.1 Has_Specialization and Has_Feature

We distinguish the relation types <Has_Specialization> (<HS>), a class-subclass relation, and <Has_Feature> (<HF>), to support inheritance. We will illustrate this with an example (see Fig. 4).

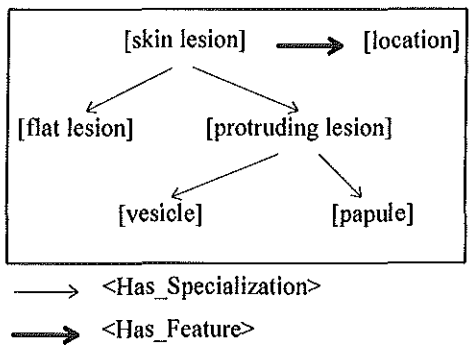


Figure 4. Contextual graph showing the relations <Has_Specialization> and <Has_Feature>. Note that [location] will only appear in the menus of [vesicle] and [papule].

In the KB, various kinds of [skin lesion] may be defined, such as a [vesicle], [papule], [tumor], etc. All these skin lesions have a location. [Vesicle] and [papule] can thus be seen as *specializations* of [skin lesion], having the *feature* location in common. When all specializations of a concept have a feature in common, then that feature is to be defined at the level of the parent.

In the entry program, features of a concept are only presented for selection to the user when that concept has no specializations. The idea is that a concept with further specializations is not specific enough to be described by features. In other words, features are *inherited* via <HS> relations, until a descendant is encountered that has no further specializations. Inheritance concerns only the presence of a feature, not its value.

When a parent concept has many specializations, or many features, and these cannot be conveniently presented to the physician, then overview is

ameliorated by introducing an intermediate level between parent and children. The concepts in such an intermediate level are then connected to the parent with <HS>. To give an example (see also Fig. 4): [Skin lesion] may be further subdivided into [protruding lesion], [flat lesion], and [excavated lesion]. These intermediate levels enable the physician to navigate conveniently through an otherwise abundant number of possible selections.

5.1.2 Preset_Choice and Exclude_Choice

When making notes in a paper record, a physician usually describes concepts from a certain perspective. Lymph nodes are described from an anatomic perspective (examination of the abdomen or examination of the neck) whereas the quality of arterial pulsations at any location are described from the perspective of cardiovascular state.

An endoscopist usually describes findings in the following order: organ (e.g., stomach) → morphology (e.g., ulcer) → specific location (e.g., fundus). This implies that when an endoscopist describes an ulcer, the possible locations are restricted to the locations of the organ he is currently describing, whereas in the KB we have defined that, in general, [ulcer] <HF> [location], where [location] can be any anatomic site. Hence, when a physician has indicated that he is describing [examination of stomach], there is a need for a mechanism that automatically selects the appropriate anatomic subdomain [stomach location].

For this purpose we have introduced the <Preset_Choice> (<PC>) relation (Fig. 5). When the selection of [A] predetermines the selection of descendant [B] , then [A] should be connected to [B] using a <PC> relation. In our endoscopy KB, [examination of stomach] is linked via <PC> to [stomach location]. The effect of a <PC> relation on the behavior of the entry program is that the sequence [examination of stomach] → [ulcer] → [location] will cause the program to immediately select [stomach location] and present its children to the user.

Analogous to the <PC> relation, it is also possible to indicate that concepts become inappropriate (other than 'all but one') after the selection of a concept. This information can be recorded using the <Exclude_Choice> (<EC>) relation.

The <PC> and <EC> relations are thus used to indicate selectional constraints that can be distilled from the path to a concept. These relations allow the entry program to adhere to the physician's usual method of

working without bothering the physician with irrelevant selections, whereas in the KB the principle of modeling general, context-free knowledge is preserved.

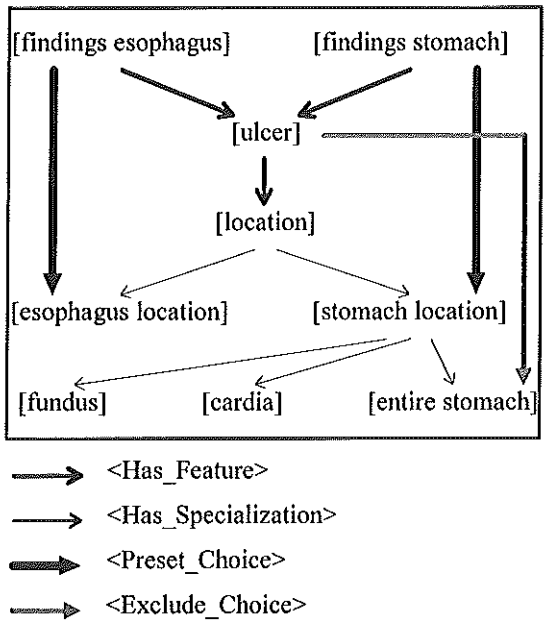


Figure 5. Contextual graph showing the use of the <Preset_Choice>, and the <Exclude_Choice> relation. When [ulcer] is selected via [findings esophagus] the entry program will only show locations in the esophagus. When [ulcer] is selected via [findings stomach] only the options [fundus] and [cardia] will be presented, while [entire stomach] is excluded.

5.1.3 Refers_To

As stated in section 4.4, the physician sequentially selects concepts when entering data. When a concept is selected for the first time, a new instance of that concept is made. When a concept is reselected, a new instance is made only when the concept path differs from the path to the existing instance of that concept. This means that when a physician has described an [ulcer] via the path including [examination of esophagus], and then reselects [ulcer], but now via [examination of stomach], a new instance of [ulcer] is made.

If a concept can only occur once in a given context, then only one

instance of that concept should be allowed to exist in that context. In the KB, this may be achieved by allowing only one path to such 'contextually singular' concepts. However, such a solution would severely limit the flexibility of the system, as physicians may want to approach contextually singular concepts from more than one perspective. For example, the physician may as well describe shortness of breath from the perspective of the history of the cardiovascular tract as from the perspective of the history of the respiratory tract, but both within the context of the patient history. If as many instances of [shortness of breath] would be created in the context of the patient history as there are paths to [shortness of breath], then the information about the shortness of breath would be scattered over more than one location in the record.

In order to keep a clear overview, to guarantee consistency, and yet to be able to offer the flexibility of multiple paths to contextually singular concepts, we introduced the <Refers_To> (<RT>) relation (Fig. 6).

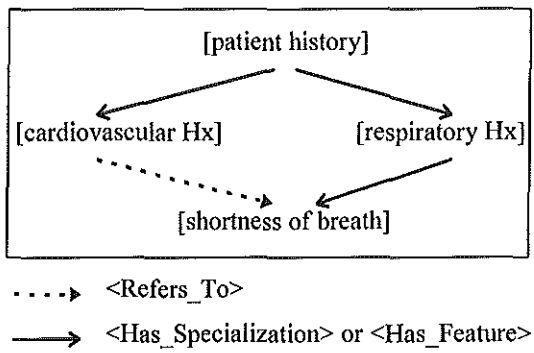


Figure 6. Contextual graph showing the use of the <Refers_To> relation. Note that the main path to [shortness of breath] leads via [respiratory Hx].

Of all multiple paths leading to a contextually singular concept, one path needs to be defined as the 'main path' in a context, meaning that contextually singular concepts may only have one parent connected through a <HS> or <HF> relation; all other parents need to be linked via <RT> relations.

In the entry program, <RT> children are presented as normally selectable options. When the concept is selected, the program will reconstruct the 'main path', and act as if the user had selected that path. The advantage of this approach is that concepts can be reached from all relevant perspectives, while all information concerning that concept is kept together. Note,

however, that [shortness of breath] via [patient history] will still lead to another instance than [shortness of breath] via [physical examination].

5.1.4 Has_Value

Some concepts are described by numeric values in combination with a unit, for example [size] 3 cm, and [weight] 70 kg. It would be impractical to connect all possible values as children to such a concept, and to present these in a menu for selection. As an alternative, we link concepts that may be quantitatively described, with Has_Value (<HV>) relations to all its possible units (Fig. 7). The concept [length], for example, has as possible units [mm], [cm], [m], etc. When a concept that has <HV> children is selected in the entry program, a special menu for value entry is presented. Using this menu, the physician can select a unit, enter a numerical value, indicate whether the value is exact or an estimate ('about 20 cm'), and construct ranges, such as '3 to 4 cm' and '5 mm by 4 cm'.

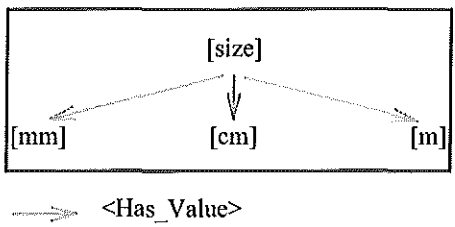


Figure 7. Graph showing <Has_Value> relations between [size] and the units by which it can be described numerically.

5.2 Properties

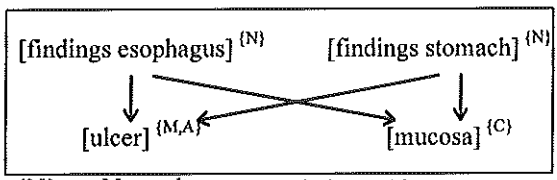
So far, we have discussed the different relations, each representing the meaning of a link between two concepts, and each with its own specific effect on the behavior of the entry program. In addition to relations, a concept may have certain intrinsic properties that characterize the use of that concept. These properties also have their own specific effect on the behavior of the entry program. The defined properties are: {multiple}, {combination}, {absent}, and {normal}; they are illustrated in Fig. 8.

5.2.1 Multiple

In section 5.1.3 we have discussed that a new instance has to be made when the selected concept path differs from the paths to existing instances of that concept. However, a new instance of [A] may also have to be made

when [A] is reselected via the same path. In case of several ulcers in the stomach, for example, it should be possible to describe each of them in detail. In the KB, we assign the property {multiple} to such concepts.

When a {multiple} concept is reselected via the same path, the entry program enables the physician to either resume describing the existing instance, or to create a new one. To provide the convenience of assigning one description to several similar instances, the user can indicate the number of instances to which that description pertains. Besides a number, the user can also express quantities with expressions such as 'several' or 'multiple'.



{N} = Normal {A} = Absent

{M} = Multiple {C} = Combination

Figure 8. Graph showing the use of the properties in the knowledge base.

5.2.2 Combination

For some concepts, it may be desirable to give more than one description, but undesirable to create more than one instance. For example, although the appearance of the mucosa may be different in various areas of the stomach, the physician will not view these as separate 'mucosas': the mucosa is still perceived as one entity. Yet, it should be possible to distinguish the following two statements:

- the mucosa is red and swollen in fundus and cardia,
- the mucosa is red in the fundus, but swollen in the cardia.

A concept receives the property {combination}, when more than one combination of features is allowed.

5.2.3 Absent

A known source of ambiguity is missing information. When a report mentions nothing about 'spider naevi', it is unclear whether or not the physician looked for them. Explicit knowledge about the presence or absence of findings is, therefore, of clinical importance. So far, we stated

that concepts, selected by the physician, represent findings present in the patient; we also need a way to indicate that an abnormality was found to be absent. Since some concepts simply cannot be absent, e.g., a concept such as 'size', we assign the property {absent} to concepts, whose instances can be absent.

If a physician states that a concept is absent, then the concept is instantiated as usual, but the finding as a whole is tagged as being absent.

5.2.4 Normal

When making notes in paper records, physicians often use abbreviations to record a set of findings at once. An example of such a shortcut is the use of the expression 'X is normal'. Although the physician using such a statement will know its explicit meaning, the expression may comprise a different meaning for other physicians, making such expressions ambiguous. The expression 'joints normal' may refer either to a careful examination of all joints with respect to pain, swelling, and degrees of movement, or to an examination in which the patient was asked to make three bendings of the knee. Not allowing a physician to use such ambiguous shortcuts would impede the acceptance of a system or lead to incompleteness. As always forcing the physician to describe in full detail what he actually examined would be considered too much effort.

Our strategy for solving this dilemma is based on the presumption that every physician has his own routine for history taking and physical examination. The explicit meaning of the expression 'normal' will thus, per physician, be fairly constant. We capture the specific meaning of each normal statement per physician and, subsequently, use this as a substitute.

In the KB, concepts for which the expression 'normal' may be used are assigned the property {normal}. When a given physician selects the option 'normal' for a concept, then the entry program checks whether that physician has already defined his description of 'normal' for the selected concept. If not, the 'normal-editor' is invoked, which functionally resembles the entry program. In the normal-editor, the physician describes which findings he wishes to include in the statement. This description will serve as a substitute each time the 'normal' option is selected again for that concept. The substitute is treated as any other description made by the physician, hence findings can be added or altered. The normal-editor can also be used independently of the entry program to modify or create definitions for normal.

Besides facilitating data entry, the advantage of interactive acquisition and substitution of physician-specific knowledge is that the patient data become independent of their author, and thus less ambiguous.

5.3 Data Entry Protocols

We offer the option of data entry protocols to promote completeness of data in a specific domain. That is, a data entry protocol dictates which concepts should be described in that domain.

To enhance uniformity in reports in which, for example, an ulcer was found, a protocol can be created defining that its size and bleeding stigmata should be described. A more general protocol may define that a proper physical examination of the lungs is described by inspection, percussion, and auscultation.

Data entry protocols are defined and handled in the entry program similar to 'normal' definitions. Protocols differ from 'normal' definitions in the sense that protocols are not physician-specific, but indicate which subjects should be described, whereas in 'normal' definitions subjects have already been described. Protocols are thus used to promote uniformity and completeness, and ease of data entry is enhanced because all subjects in the protocol definition are directly selectable for further description.

5.4 Free Text

Theoretically, any information in medical records may be structured, even Rector's example of a patient's chest pain occurring when walking past the freezing compartment of a supermarket. But, as this information is more anecdotal than medically relevant [13], structuring such information would not be meaningful. However, as the physician should never feel limited in his expressive power, we give the physician the option to add free text to any instance. These free-text additions will enable us to study the limitations of SDE: where structuring of data is impossible, and where free text is preferred even when structured options exist.

6. Discussion

In this paper, we have focused on *descriptional knowledge*. Making this type of knowledge *explicit* is needed for flexible structured *entry* of descriptive data such as findings in the patient's history, physical examination or additional examinations. Once such data can be captured in

a structured form, there is a potential for more reliable interpretation and use of these data for patient care and a variety of other purposes. We will now explain how the model meets the requirements for SDE, which we formulated in section 2.

With regard to the *expressive power* of the model we conclude that when there are multiple occurrences of an abnormality, each can be described in detail, but general statements about multiple occurrences can also be made. Descriptions of fine granularity can also be achieved by using combinations of descriptors to specify multiple appearances of one abnormality. Furthermore, the model has no limitations with respect to the depth or breadth in which topics may be described, thus allowing theoretical domain completeness.

The model provides mechanisms for *flexible* data entry that *corresponds to the physician's usual method of working*. The physician has the freedom to determine the order and depth in which he describes findings. Any concept can be directly accessed at any point, by using a search keyword. Therefore, it is not required that the physician knows in advance how to navigate to the concept he wishes to select. Where appropriate, concepts can also be accessed from different perspectives, as different physicians may desire. Yet, overview is maintained, as entered data retain the order defined in the knowledgebase: data entered by colleagues have the same *predictable order* as the data entered by oneself. Furthermore, the use of summarizing terms such as 'normal' is supported.

SDE should *cost no more time than free-text reporting*, although it generally does. However, we believe that in case of the (more) frequently occurring normal findings the use of the 'normal' option and the use of protocols will, in practice, save time. Thus, although the explicit description of abnormalities will be more labor intensive, we feel that overall, SDE will be time efficient.

More complete data is promoted through the reminder effect of showing all possible descriptors of a concept in its context. Physicians may, furthermore, be cued to what should be described by data entry protocols.

We *reduce ambiguity* of obtained data in several ways. First of all, as statements such as 'X is normal' may have different meanings for different physicians, these statements are replaced by their explicit meaning during data entry. Second, the model allows the physician to record the absence of findings, thereby enabling a distinction with findings that were not examined. Third, as the context of instances of concepts is uniquely

determined, the patient data are stored unambiguously. Therefore, concepts in the descriptional knowledgebase do not necessarily need to have exactly one meaning, in contrast to what Cimino states [24], thus avoiding the problem of knowing where to stop disambiguating concepts (e.g., is 'ventricular ulcer' unambiguous, or should 'ulcer in the cardia' and 'ulcer in the fundus' be distinguished).

We emphasize that the model presented in this paper should not be confused with medical concept representations, as described in [13,18,19,31-35]. In our approach, the representation of the eventually obtained patient data (findings) is consonant with other proposed schemata such as conceptual graphs [30-32]. Descriptional knowledge, however, does not deal with the representation of medical data in general, but deals with representing the knowledge needed for flexible and expressive data entry. As descriptional knowledge should be seen as an extension of currently available medical concept representations, we believe that descriptional knowledge could easily be added to medical concept representation schemes.

7. Current Status and Research Agenda

We have implemented our contextual graphs model as a stand-alone prototype (Fig. 9). Using the knowledge-base editor, two different knowledgebases are currently being developed: one for use in an internal medicine outpatient clinic, and one for the endoscopy domain.

Our model, however, still requires extensions; at present we are not able to indicate (un)certainty of observations, to use synonymy, or to perform additional validity checks. We are extending the idea of capturing physician-specific knowledge: in analogy with 'normal' definitions, physicians will be able to define their own summarizing terms, such as 'gastroenteritis complaints'. In addition, we need to develop an extra level of expressiveness to indicate relations between findings (that is between instances) to specify, for example, that a patient's headache *co-occurred* with nausea, or that an ulcer was located *in* or *just above* a diverticulum.

At present, the possible relations between and the intrinsic properties of concepts are the subject of our research; so far, we have not used existing standard vocabularies such as SNOMED III [14] as source for our thesauri. Compatibility or integration with existing standards will, therefore, be another important topic for future research.

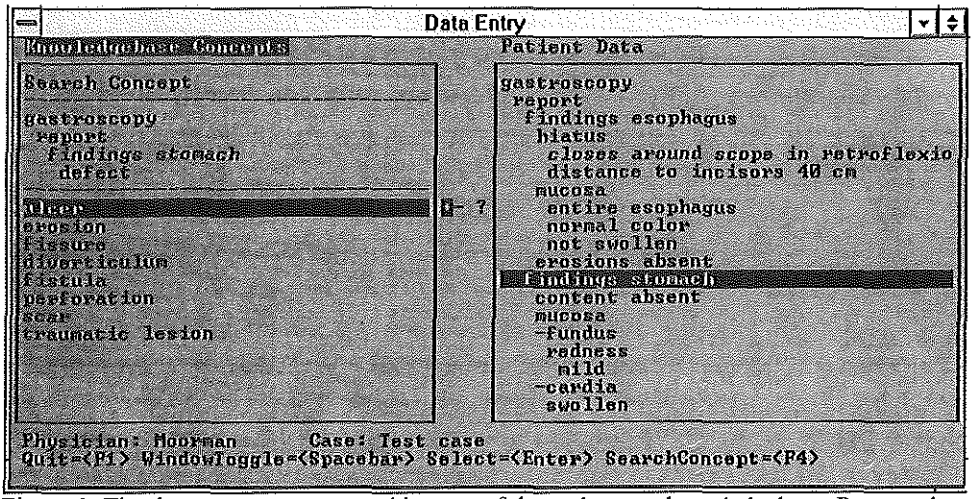


Figure 9. The data entry program, making use of the endoscopy knowledgebase. Presented on the left side are the concepts as present in the knowledgebase. The concepts between the dashed lines are in the concept path. The + and the - are used to indicate whether a concept is present or absent. The ? is intended for later use. On the right side the (instances of the) concepts selected by the physician are presented.

An evaluation of the system using the endoscopy KB is the next step scheduled. Central research issues of this evaluation will be the expressive power of the model, and the resulting degree of unambiguity and completeness of the data recorded in the reports.

Acknowledgment

The authors thank Dr. Mark Musen, Stanford University, USA, for his valuable contribution to this paper. This study was financially supported by Glaxo BV.

References

1. Dick RS, and Steen EB, eds. *The Computer-Based Patient Record. An Essential Technology for Health Care.* Washington DC: Institute of Medicine, 1991: 15-9.
2. Hammond WE, and Stead WW. The evolution of a computerized medical information system. In: Orthner HF, ed. *Tenth SCAMC.* New York: IEEE Computer Society Press, 1986: 147-56.
3. Stead WW, and Hammond WE. Computer-based medical records: the centerpiece of TMR. MD Comput 1988; 5: 48-62.

4. Barnett GO. The application of computer based medical record systems in ambulatory care. New Engl J Med 1984; 310: 1643-50.
5. McDonald CJ, Tierney WM, Overhage JM, et al. The Regenstrief medical record system: 20 years of experience in hospitals, clinics, and neighborhood health centers. MD Comput 1992; 9: 206-15.
6. Van der Lei J, Duisterhout JS, Westerhof HP, et al. The introduction of computer-based patient records in the Netherlands. Ann Intern Med 1993: 119: 1036-41.
7. Young DW. What makes doctors use computers? Discussion paper. In: Anderson JG, Jay SJ, eds. *Use and Impact of Computers in Clinical Medicine.* New York: Springer-Verlag, 1987: 8-14.
8. Safran C, Porter D, Lightfoot J, et al. ClinQuery: A system for online searching of data in a teaching hospital. Ann Intern Med 1989; 111: 751-6.
9. Kent DL, Shortliffe EH, Carlson RW, et al. Improvement in data collection through physician use of a computer-based chemotherapy treatment consultant. J Clin Oncol 1985; 3: 1409-17.
10. Gouveia-Oliveira A, Raposo VD, Salgado NC, et al. Longitudinal comparative study on the influence of computers on reporting of clinical data. Endoscopy 1991; 23: 334-7.
11. Kaplan B. The influence of medical values and practices on medical computer applications. In: Anderson JG, Jay SJ, eds. *Use and Impact of Computers in Clinical Medicine.* New York: Springer-Verlag, 1987: 39-50.
12. Reiser SJ. The clinical record in medicine. Part 2: reforming content and purpose. Ann Intern Med 1991; 114: 980-5.
13. Lindberg DAB, Humphreys BL, and McCray AT. The Unified Medical Language System. Meth Inform Med 1993; 32: 281-91.
14. Rothwell DJ, Côté RA, Cordeau JP, and Boisvert MA. Developing a standard data structure for medical language - the SNOMED proposal. In: Safran C, ed. *Seventeenth SCAMC.* Ney York: McGraw-Hill, 1993: 695-9.
15. Rector AL, Nowlan WA, and Kay S. Foundations for an electronic medical record. Meth Inf Med 1991; 30: 179-86.
16. Rector AL, and Kay S. Descriptive models for medical records and data interchange. In: Barber B, ed. *Medinfo 89.* Amsterdam: Elsevier Science Publishers, 1989: 230-4.
17. Rector AL, Nowlan WA, and Kay S. Unifying medical information

using an architecture based on descriptions. In: Miller RA, ed. *Fourteenth SCAMC*. New York: IEEE Computer Society Press, 1990: 190-4.

18. Rector AL, Nowlan WA, Kay S, et al. A framework for modelling the electronic medical record. Meth Inf Med 1993; 32: 109-19.
19. The Galen Project. Project 2012, Aim Initiative. Deliverable 6. Manchester, UK, 1993.
20. Nowlan WA, and Rector AL. Medical knowledge representation and predictive data entry. In: Fieschi M, Talmon J, eds. *Proceedings of AIME-91*. Heidelberg: Springer Verlag, 1991: 105-16.
21. Trace D, Naeymi-Rad F, Haines D, et al. Intelligent medical record - entry (IMR-E) J Med Sys 1993; 17: 139-51.
22. Campbell KE, and Musen MA. Creation of a systematic domain for medical care: the need for a comprehensive patient-description vocabulary. In: Lun KC, ed. *Medinfo 92*. Amsterdam: Elsevier Science Publishers 1992: 1437-42.
23. Campbell KE, Wieckert K, Fagan LM, and Musen MA. A computer-based tool for generation of progress notes. In: Safran C, ed. *Seventeenth SCAMC*. New York: McGraw-Hill, 1993: 284-8.
24. Cimino JJ, Hripcsak G, Johnson SB, and Clayton PD. Designing an introspective, multipurpose controlled medical vocabulary. In: Kingsland LC, ed. *Thirteenth SCAMC*. New York: IEEE Computer Society Press 1989: 513-8.
25. Moorman PW, Van Ginneken AM, Van der Lei J, et al. The contents of free-text endoscopy reports: An inventory and evaluation by peers. Endoscopy 1994; 26: 531-8.
26. Bell DS, Greenes RA, and Doubilet P. Form-based clinical input from a structured vocabulary: Initial application in ultrasound reporting. In: Frisse ME, ed. *Sixteenth SCAMC*. New York: McGraw-Hill, 1992: 789-90.
27. CEN TC251/PT003. *Medical Informatics - Structure of Concept Systems - Vocabulary* (Project Team Mose). Brussels: CEN, 1993.
28. ISO 2788. *Documentation - Guidelines for the Establishment and Development of Monolingual Thesauri.* Geneva: International Organization for Standardization, 1987.
29. ISO 1087. *Terminology - Vocabulary.* Geneva: International Organization for Standardization, 1990.
30. Sowa JF. *Conceptual Structures*. Reading MA: Addison Wesley, 1984.

31. Campbell KE, and Musen MA. Representation of clinical data using SNOMED III and conceptual graphs. In: Frisse ME, ed. *Sixteenth SCAMC*. New York: McGraw-Hill 1992: 354-8.
32. Campbell KE, Das AK, and Musen MA. A logical foundation for representation of clinical data. J Am Med Inform Assoc 1994; 3: 218-32.
33. Cimino J, Clayton PD, Hripcsak G, and Johnson SB. Knowledge-based approaches to the maintenance of a large controlled medical terminology. J Am Med Inform Assoc 1994; 1: 35-50.
34. Evans DA, Cimino JJ, Hersch WR, et al. Toward a medical-concept representation language. J Am Med Inform Assoc 1994; 1: 207-17.
35. Friedman C, Cimino JJ, and Johnson SB. A schema for representing medical language applied to clinical radiology. J Am Med Inform Assoc 1994; 1: 233-48.

Chapter 7

Evaluation of reporting based on descriptional knowledge

Moorman PW [a,b]
van Ginneken AM [a]
Siersema PD [b]
van der Lei J [a]
and van Bemmel JH [a]

[a] Department of Medical Informatics, Erasmus University,
[b] Department of Internal Medicine, University Hospital Dijkzigt,
Rotterdam, The Netherlands

J Am Med Inform Assoc (in press)

Abstract

In our attempt to enhance the completeness and clarity of clinical narratives, we developed a general formalism for the entry of structured data. The objective of this study was to gain insight in the expressive power of the formalism, and the effects of its use on reporting in endoscopy.

Ten endoscopists reported twice on 8 endoscopy videotapes: they first produced free-text reports and then structured reports using our formalism. Statements in the resulting reports were compared for both methods.

In total, 6.8% of the endoscopists' statements could not be expressed in structured options. Most of these statements were not due to limitations of the formalism itself. Most topics that were described in free-text reports were more often mentioned in structured reports. Furthermore, structured reports addressed topics that had not been described in free-text reports. Overall, we observed an increase of 83% for topics not related to abnormal findings, and 45% for features of abnormal findings. Although there was an overall information gain, features of abnormal findings were on average still described only by half of the endoscopists.

The expressive power of our formalism is promising, but general, multipurpose, usage of the acquired data, requires that topics are described by a larger percentage of physicians. Since our formalism led to more complete and more uniform data, additional research is justified to study how spontaneous reporting can be augmented further. The few subjects that occurred less often in structured reports suggest a possible negligence effect of structured reporting.

Introduction

Because data in paper-based patient records have limited suitability for formal analysis, computer-based patient records (CPRs) increasingly gain interest [1]. In early CPRs coding was mainly confined to laboratory data, medications and diagnoses, whereas the narratives, such as findings of history taking and physical examination, were recorded in free text [2-6]. Free-text data, even in electronic form, have drawbacks of which spelling errors, ambiguity and incompleteness are the major ones. Although many efforts are undertaken to obtain coded data from free text, techniques such as natural language processing do not improve the quality of the recorded data. Formalisms that a priori try to collect data in a structured, coded format are more likely to increase the usefulness of the data for research, decision support, quality assessment, and clinical care itself.

When the domain of the data that are to be captured, is small and well circumscribed the use of paper- [7] or computer-based [8-13] forms has shown to be feasible, particularly when the data elements that need to be captured are well defined. However, when the domain becomes large, forms become impractical to accommodate the flexibility to which the physician is used with free text [14]. Furthermore, limiting data capture to essential elements restricts the expressive power.

To overcome these limitations we developed a formalism based on explicit descriptional knowledge [15]. Descriptional knowledge is general knowledge that describes where, when and how concepts can be described. The aim of our formalism was to enable the capture of structured, coded data with an expressive power approaching that of free text, while maintaining flexibility and reducing ambiguity. To evaluate these objectives, we built a data-acquisition front-end for data entry conforming to our formalism: the *entry-program*. The behavior of this general entry-program is determined by a combination of user input and domain-specific descriptional knowledge. This domain-specific descriptional knowledge is stored in a knowledgebase. We developed knowledgebases for general internal medicine (physical examination), gastrointestinal endoscopy, pathology (cutaneous lymphomas), and radiology (chest X-rays).

For formalisms aimed at capturing clinical data a prime criterion for acceptability is the ability to express the clinical findings. In this study we focus on the expressive power of a structured data entry formalism in a

specific domain. The purpose of the study is not to assess the practical implication and feasibility of using structured data entry in a routine setting, nor to demonstrate that a change of report practice results in better care. This study attempts to explore the limitations of our structured data entry formalism. We sought to answer the following questions:

- how good is the expressive power of our formalism? Can physicians express what they want when restricted to our predefined structure and terminology?
- what happens to the quantity of the acquired data? Using structured data-entry, other researchers [11-13] have shown that the 'completeness' of small sets of items increased compared to free-text reports. Is such an increase also feasible when the items that can be captured extend beyond such limited sets?
- what happens to the uniformity of the acquired data? Do reports by multiple observers, describing the same examination and produced with our formalism, resemble each other more than free-text reports resemble one another?

We chose to study the domain of endoscopy; a previous study showed that the size and complexity of this domain is large enough to require an approach other than forms [16]. That study furthermore showed that endoscopists themselves indicate that their currently produced free-text reports do not contain sufficient detail.

Methods and Materials

Study design, participants and study material

For this evaluation, we have chosen an experimental setting in which esophago-gastro-duodenoscopy video tapes were described by 10 endoscopists, both in free text and with the entry-program. Two endoscopists work in our university hospital, the other 8 in hospitals affiliated with our university hospital.

We videotaped the complete examination of patients undergoing an esophago-gastro-duodenoscopy. The first 8 videos that contained at least one abnormality were selected. In total the 8 videos showed 14 abnormal findings.

For each video, each endoscopist prepared a free-text report in his usual fashion (9 dictated the reports, and 1 wrote them). Two to four weeks later,

the endoscopists prepared the reports using the entry-program, after the videos had been shown again. The endoscopists had never used or seen the program before, and were given a short introduction on the use of the program (about 10 minutes). A note was made of any observation that could not be expressed in structured options.

Throughout the study no constraints were placed on the amount of time the physician wanted to spend on recording the findings.

Short description of the entry-program

The entry-program is menu driven, and its behavior is determined by a combination of user input and the content of the used descriptional knowledgebase. The endoscopy knowledgebase basically consists of concepts, each of which exists only once in the knowledgebase. Concepts have relations with other concepts, and the entry-program uses these relations to show for each selected concept its descriptors: all concepts describing that concept in a given context. This process is repeated after a selection by the user. Furthermore, in the knowledgebase is represented whether concepts may occur multiple times in a context, and whether a physician may state that the concept is absent. In this thesis, the complete contents of the gastroscopy knowledgebase is listed in the Appendix. During data-entry, the physician may also use 'normal definitions' in our formalism: each physician only needs to define once what he or she means when using a statement such as 'findings in the esophagus normal'. When used in reporting, the explicit, physician-specific meaning of such a statement is then incorporated in the report. For a detailed description of the formalism, see [15].

We will use the term *structured report* when we refer to reports made with the entry-program.

Three types of data cannot be expressed with the current prototype and knowledge-base: uncertainties, performed actions (e.g., biopsies) and relations between described findings (e.g., it can be stated that there is an ulcer and a polyp, but it is not possible to formally state that the ulcer was located *on* or *near* the polyp).

Analysis of reports

In this study we restricted ourselves to the description of findings; other components of reports, such as indications and conclusions, were not included.

Free-text reports. To enable comparison of reports, we first made an inventory of the contents of the free-text reports by identifying all *statements*. A statement is defined as each combination of a *subject* and *described feature*. In Table 1 we provide an example of a free-text report and its statements.

Table 1. Example of a free-text report and its statements

	Statements	
Free-text	Subject	Described feature
The esophagus is covered with normal mucosa.	Esophagus mucosa	Aspect
The Z-line is situated directly above the hiatus.	Z-line and hiatal impression of diaphragm	Position with respect to each other
In inversion, we see that the hiatus closes around the scope.	Hiatus	Closure around scope
The gastric mucosa is intact everywhere.	Gastric mucosa	Aspect
The pylorus has an oval shape, and can easily be passed.	Pylorus	Shape Passage
In the anterior wall of the bulb, we see a deep, more or less regular*, round-shaped ulcer, covered with fibrous exudate. The border is regular. The diameter is 5 to 6 mm.	Ulcer	Location site Anatomic location Depth Base/shape *- regularity Shape Type/color of exudate Border - regularity Numerical size
In the descending duodenum intact mucosa	Descending duodenum mucosa	Aspect

* From the free-text description it remains unclear to what feature 'regular' refers.

All statements from the free-text reports were assigned to one of the following categories:

- *General statements.* These are statements that appeared in each group of reports. The term 'group' refers to all reports describing the same video. Hence, general statements are abnormality independent. Examples of these statements are: the position of the Z-line with respect to the diaphragm impression, and the shape of the pylorus.
- *Features.* For each of the 14 abnormal findings, a list of features describing that finding was made. Examples of these features are the size and shape of an ulcer.
- *Unclear feature descriptions.* Features were placed in this category when it was unclear to which feature a description referred. In the

example of Table 1, the expression 'irregular ulcer' is considered unclear as it may as well refer to the regularity of the base as to the regularity of the shape.
- *Other statements.* Any statement not falling in one of the above 3 categories.

In addition, we constituted per abnormal finding a list with the *type-labels* that the endoscopists used to name an abnormal finding. For example, one endoscopist may use the term 'erosion' while another endoscopist would use 'ulcer' to describe the same lesion.

Structured reports. The same inventory of statements and lists was made for the contents of structured reports. An additional list was made containing all statements that endoscopists could not express using structured options in the entry-program.

Comparison of reports. We compared the following topics:
- General statements: number of different general statements, and number of times that they were mentioned.
- Features: number of different features per abnormal finding, and number of endoscopists that mentioned each finding.
- Other statements: total number of mentioned statements in this category.
- Type-labels: the number of different type-labels per abnormal finding.

To gain insight into the nature of the unclear statements, we asked the endoscopists to attempt to express their unclear statements in our formalism, after they had made their structured reports.

Results

Quantity of the data

In total, the free-text reports contained 871 statements: 366 general statements, 406 features, 60 other, and 39 unclear statements. The structured reports contained 1297 statements: 671 general statements, 586 features, and 40 other. In total, 88 statements could not be expressed with the entry-program.

The occurrence of general statements and features in the reports is now

discussed, and summarized in Table 2.

Table 2. Number of different general statements and features, and number of times they were described in free-text and structured reports.

	Nr of different statements	Nr of times described	
		in free-text reports	in structured reports
General statements			
in free-text and structured reports	8	366	448
new in structured reports	11	-	223
Features			
in free-text and structured reports	101	406	526
new in structured reports	22	-	64

General statements

In the free-text reports we identified 8 different general statements: these were mentioned 366 times in the free-text reports. These same 8 general statements were mentioned 448 times in the structured reports. In the structured reports, 11 additional statements fell in the category of general statements. These were mentioned 223 times. Thus, the total number of mentioned general statements increased 83% from 366 in the free-text reports to 671 in the structured reports.

Features

In the 8 videos, 14 abnormalities were present. In free text, an abnormality was on average described by 7.2 features (in total 101 features), and every mentioned feature was on average described by 4.0 endoscopists (in total 406 described features). These same features were described by, on average 5.2 endoscopists in structured reports (an increase of 30%). Of the 101 free-text features, 22 were described by as many endoscopists in structured reports, 25 were described by fewer endoscopists in structured reports, and 54 were described by more endoscopists in structured reports. Furthermore, on average, 1.6 more features per abnormality were described in structured reports (in total 22 more features, increase of 22%), which were on average described by 2.8 endoscopists. Overall, the 123 structured features were on average described by 4.8 endoscopists (in total 590 described features).

Of the statements that were described less often in structured reports, the most striking decrease concerned a video in which esophagus varices were present. In free text, 8 endoscopists made the statement 'varices in the

fundus are not present' (i.e. no varices in the stomach), while only one endoscopist made this statement in the structured reports.

Expressive power of the model and knowledge-base

During data entry, the endoscopists felt that 88 statements could not be expressed in structured options. As a statement was sometimes made by different endoscopists, or made by the same endoscopist in different reports, 51 different statements were identified. As shown in Table 3, we grouped these statements according to the modification that would be required to allow expression of these statements in structured options.

Table 3. Statements that endoscopists could not make in structured options, grouped according to required modification

	number of different statements	number of times mentioned
Knowledge-base modification	46	77
Concept change	4	5
Addition of relations	13	30
Addition of concepts and relations	8	10
Debatable	21	32
Model adjustment	5	11
Total	51	88

Modification of the knowledgebase: 46 of the 51 different statements.

Four of the statements would be resolved by changing a concept. We give two examples. The term 'circular', in the knowledgebase used to denote a 'shape', should be changed to 'round', as some endoscopists preserved the term 'circular' to denote 'covering the complete circumference of (e.g.) the esophagus'. The concept 'signs of previous bleeding' should be changed to 'signs of recent bleeding'.

Thirteen of the statements would be resolved by adding relations between existing concepts in the knowledgebase. For example, some endoscopists wanted to describe the 'surrounding mucosa' of a 'scar', the 'peristalsis' in the duodenum, and the 'part' (e.g., 'distal', 'middle', and 'proximal') of the location 'bulb'.

Eight of the statements would require the introduction of new concepts. For example, two endoscopists wanted to describe a shape using the term 'triangular', and preferred to use 'hours' to describe the site of findings in tubular organs instead of the options 'left', 'right' 'front' and 'back'.

Twenty-one statements would require careful consideration before

adding these to the knowledgebase: they would possibly introduce redundancy, ambiguity, or would add no 'relevant' detail. Two types of statements that introduce redundancy can be distinguished: the statement is inherent to a concept, or the statement can be formulated using concepts in the knowledge-base. An example of the first type is that some endoscopists wanted to state that an erosion was 'superficial', or was 'red'. However, by definition an erosion is superficial, and when an erosion is not covered by exudate (which is one of the options by which to describe an erosion) it is always red. It is therefore unclear whether such statements add 'relevant' detail. An example of the second type is that endoscopists were looking for words such as 'particularly', or 'most pronounced', e.g., to state that the mucosa was red in the fundus and antrum, but most pronounced in the antrum. Although it would be possible to add the possibility of such statements to the knowledgebase, we feel that the endoscopists probably meant something as 'there is mild redness in the fundus, and severe redness in the antrum'. Therefore, addition of concepts as 'particularly' would not only introduce redundancy, but also ambiguity.

Model adjustment: 5 of the 51 different statements.

Five of the statements would require an extension of the model. Although all fall into the category 'relations between findings' e.g., the pylorus is asymmetrically deformed *by* an ulcer in the bulb, it will also be necessary to support statements about the arrangement of multiple occurring identical abnormalities: e.g., erosions occur throughout the stomach, but lie around in groups or in rows.

Uniformity and ambiguity.

In free text, we found on average 5.0 different type-labels to name a given abnormal finding; in structured reports this had decreased to 2.2. Whereas in free text, none of the abnormalities were named with the same type-label by all endoscopists, in structured reports all 10 endoscopists used the same type-label for 5 of the 14 abnormalities.

In the free-text descriptions of the abnormal findings, we encountered 39 unclear feature descriptions. When asked to do so, the endoscopists had no trouble in translating those statements into knowledge-base concepts. All unclear feature descriptions that we encountered in the free-text descriptions of *ulcers*, together with their knowledge-base translations are listed in Table 4.

Table 4. Unclear descriptions of ulcer features, and their translation into knowledgebase concepts.

Expression in free text	Translated in knowledge-base concepts
Ulcer is 'punched out'	Ulcer is deep, sharply demarcated, and surrounding mucosa has normal color
Exudate of ulcer is elevated	Border is elevated
Ulcer in local thickening	Border is elevated
Base is messy	Base is irregular
Base is clean	Base contains white exudate
Base is smooth	Base is regular
Irregular ulcer	Shape and base are irregular
Border is quiet	Border is regular
Surrounding mucosa is quiet	Surrounding mucosa has normal color, and is not swollen
Border is sharp	Ulcer is sharply demarcated
Border is red and swollen	Surrounding mucosa is red and swollen

Discussion

The main objective of this study was to gain insight in the differences between reports expressed in free text and reports expressed in our formalism. For this purpose, 10 endoscopists reported on 8 endoscopy videotapes with both methods, and we studied the expressive limitations of our formalism, the quantity of the data, and assessed differences in uniformity.

Although several evaluations of programs for structured data entry (SDE) are described in literature, only a few describe more than user-acceptance and/or time requirements alone. The main focus of our research is the expressiveness of our SDE formalism; we did not attempt to assess practical feasibility or time requirements.

Kuhn [11], Bell [12], and Gouveia-Oliveira [13] have studied 'completeness' of free-text and structured reports in the domains of upper abdominal ultrasound, pelvic ultrasound, and endoscopy, respectively. These three studies differ from our study with respect to study design. The three studies evaluated reports produced in a clinical setting, whereas our study was done in an experimental setting. This allowed us to study reports describing the same examination.

Furthermore, in assessing 'completeness' of reports, the studies took as starting point 'elements' considered essential and/or of great clinical importance. In other words, they were restricted to elements that always need to be described in a report, or in the description of a given finding.

This is what we refer to as the form-based approach. However, in a previous inventory [16], we found that in endoscopy reporting, such 'essential' elements hardly exist, and that the presence of a description of an element depends on the circumstances in which it is found. Including the complete descriptive contents of reports in our evaluation allowed us insight beyond essential elements alone.

Moreover, how often a statement is regarded as present will depend on its studied detail. Gouveia-Oliveira, for example, studied whether the border of an ulcer was described, whereas we studied whether the regularity, elevation etc. of the border of an ulcer was described.

Finally, Kuhn [11], and Bell [12] took the percentage of structured reports with free-text annotations as a parameter for expressiveness. We feel that such a parameter is less informative as it gives no insight in the proportion of information formulated in free-text.

In our experimental setting endoscopists may have been more motivated to describe features than in a clinical setting. This is, for example, illustrated by the fact that in the free-text reports of this study the numerical size of abnormalities was described in almost 90% of the cases, whereas it was only described in 30% in an earlier report inventory [16]. However, as circumstances for both methods of reporting were equal in our study, over-motivation will hardly have affected comparison of both methods.

With regard to *expressive power* of our formalism, we conclude that relatively few statements could not be expressed. If we compare the 88 statements that could not be made to the 1297 statements that the endoscopists made in their 80 structured reports, then 6.8% of their statements could not be expressed with the current model and endoscopy knowledgebase. Moreover, half of these 88 statements would require only minor additions to, or modifications of the knowledgebase.

One-third of the 88 statements remain debatable as to whether they should be added as structured options. Reasons for not adding those to the knowledgebase would be that they possibly would introduce unnecessary ambiguity or redundancy, or that they would add no 'relevant' detail. That this will require extensive debate is shown by the following example. From the knowledge-base developers point of view, adding concepts to describe 'depth' to the descriptors of erosion would not add any information, as an erosion is by definition superficial. It may be that the endoscopist, who wanted to make such a statement, would agree that it was indeed self-

explanatory. However, the endoscopist possibly wanted to add 'superficial' on purpose, to accentuate this for the physician who referred the patient, and who may not have sufficient endoscopic knowledge to appreciate the terms used.

As the five statements that would require extension of the model all dealt with 'relations between findings', adding such expressiveness to our formalism has high priority.

With regard to the *quantity of collected data*, we first want to discuss whether it is a relevant parameter. In a previous study [16] we showed that endoscopists are dissatisfied with currently produced free-text reports, and concluded that they do not report according to what they think should be reported. This finding led to our assumption that any increase in quantity of the reported data should be regarded as a gain. The topic of whether a newly described item has clinical significance is virtually impossible to answer, particularly since the previous study also showed that consensus among endoscopists was nearly absent on the question whether an item should be included in a given report. Therefore we took as starting point that what an endoscopist wants to state must be possible to state, as long as it does not cause redundancy or ambiguity.

In the discussion of change in the *quantity of collected data*, we distinguish the general statements from the statements on features of findings.

An increase of 83% was seen in the documentation of general statements. Particularly the increase in the number of different general statements (8 vs. 19) accounted for this increase. The reason for this increase probably lies in the substitutions of 'normal definitions' in our formalism.

With regard to features, we noted that features that had been described in free-text reports were described 30% more often with our formalism. Furthermore, there was an increase of 22% in the number of described features per abnormal finding.

As our comparison extended beyond 'essential elements', we were able to demonstrate that also new elements showed up in structured reports, and that the occurrence of elements in structured reports was associated with the a priori occurrence in free-text reports. For example, new features were on average described by only 2.8 endoscopists, whereas the features that had already been present in free-text reports were described by 5.2 endoscopists.

A peculiar observation was that in the reports concerning the video in which esophageal varices were described, the absence of fundus varices was no longer reported in the structured reports. In our descriptional knowledgebase, the option to describe fundus varices is not offered when esophageal varices have been selected. Thus fundus varices are not brought to the attention when describing esophageal varices. The low description rate of the fundus varices may be explained by the fact that the endoscopists were so pre-occupied with the offered options for description, that they forgot their 'normal clinical thinking'. Of course, this may be a transient issue, which wears off when endoscopists get more experienced in using structured reporting, but it may also point to a potential danger of structured reporting: topics that are not brought to the attention, may also be neglected. This observation stresses the fact that programs using SDE should be 'as complete as possible' as SDE may not only have a reminder effect, but also a negligence effect. In our descriptional knowledge model a solution is already provided: to a menu, options can be added that do not really describe the preceding concepts, but are cross-references to other concepts.

Furthermore, we noted that 25 features were described by fewer endoscopists in the structured reports. These features were so miscellaneous that a structural mechanism is difficult to pinpoint. It may be that, because physicians could not express uncertainties, they preferred not to describe a feature of which they were not completely convinced.

Inherent to the use of structured reporting is that reports become more *uniform*. In this study this was shown by a) an increase in number and description rate of general statements, b) an increase in the number of endoscopists describing a feature, c) an decrease in the number of type-labels used to name an abnormal finding, and d) the fact that unclear statements could be translated into knowledge-base statements.

Despite these promising results, conclusions with regard to the suitability of the acquired data for general, multi-purpose use are less positive. For formal analysis, it would be ideal when endoscopists would use the same type-labels to name abnormal findings, and would describe the same topics for a given examination. However, in this study all endoscopists used the same type-label for only 5 of the 14 abnormal findings, and features were on average only described by 4.8 endoscopists. Furthermore, one has to

consider that we took the fact that a feature *was* described as parameter for uniformity, and not *how* it was described (e.g., whether all endoscopists use the same option, when they describe border-regularity). Such interobserver variability was the subject of a previous study [17].

Limitations

We performed this study for the endoscopy domain, and in an experimental setting. Although our formalism is general and developed to enable the capture of many types of medical descriptive information, it remains to be seen whether the results of this study will also apply in other medical domains. Therefore, we tried as explicitly as possible to distinguish, in the discussion of the results, model shortcomings from shortcomings of the specifically studied endoscopy knowledgebase. Although the model cannot yet represent complex temporal relationships, the model looks also promising in other domains involving momentary observations, as we have experienced in the domains mentioned in the introduction.

In this study we did not attempt to mimic a clinical environment, and the results are therefore difficult to generalize to settings where clinicians are under time-pressures. However, we considered an experimental setting more suitable to study the step preceding practical feasibility, namely expressive power. When endoscopists would have refrained from describing certain findings due to time pressure, we would not have gained insight as to whether those findings could have been expressed. If the experimental setting would not have yielded a gain in reported facts, the need for an evaluation under clinical time pressure would have been eliminated.

Furthermore, it may be argued that more reporting is not necessarily better reporting. Although endoscopists indicate that more detailed reporting is useful, the study does not provide a clinical evaluation of the generally shared opinion. Nor does the study take into account that more reported facts may decrease the subsequent readability of a report.

Conclusions and Recommendations

We conclude that our formalism offers a promising expressive power. The underlying descriptional knowledge model first of all needs to be

extended with the possibility to represent relations between findings to further enlarge expressive power.

We demonstrated that with structured reporting the quantity of recorded data increased. The amount of increase, however, depended on the type and the a priori presence of data in free-text reports. In structured reports, subjects were not only described more often, but also new subjects were described. This indicates that an increase of data is also possible beyond the form-based, 'essential-elements-only' approach.

Despite the increase in 'completeness' and uniformity in structured reports, we conclude that the suitability of the acquired data for general, multi-purpose, usage is slightly sub-optimal. Recent studies, however, indicate that physicians may record more data in a formal way, once it is to their benefit [18, 19]. Furthermore, it should be noted that the participants in this study were all experienced physicians. Therefore, we believe that, besides additional research to improve the quantity of spontaneously reported data, it will be valuable to study the educational effects of formal reporting by physicians who are just starting the reporting routine. Finally, further evaluations of our formalism are needed to assess whether the results of this study also apply to other domains.

Acknowledgment

The authors thank the endoscopists for their kind participation. This study was financially supported by Glaxo BV.

References

1. Dick RS, and Steen EB, eds. *The Computer-Based Patient Record. An Essential Technology for Health Care.* Washington DC: Institute of Medicine, 1991: 15-9.
2. Hammond WE, and Stead WW. The evolution of a computerized medical information system. In: Orthner HF, ed. *Tenth SCAMC.* New York: IEEE Computer Society Press, 1986: 147-56.
3. Stead WW, and Hammond WE. Computer-based medical records: the centerpiece of TMR. MD Comput 1988; 5: 48-62.
4. Barnett GO. The application of computer based medical record systems in ambulatory care. New Engl J Med 1984; 310: 1643-50.
5. McDonald CJ, Tierney WM, Overhage JM, et al. The Regenstrief

medical record system: 20 years of experience in hospitals, clinics, and neighborhood health centers. MD Comput 1992; 9: 206-15.
6. Van der Lei J, Duisterhout JS, Westerhof HP, et al. The introduction of computer-based patient records in the Netherlands. Ann Intern Med 1993; 119: 1036-41.
7. Wrenn K, Rodewald L, Lumb E, et al. The use of structured, complaint-specific patient encounter forms in the emergency department. Ann Emerg Med 1993; 22: 805-12.
8. McDonald CJ, Tierney W, and Blevins L. The benefits of automated medical record systems for ambulatory care. In: Orthner HF, ed. *Tenth SCAMC*. New York: IEEE Computer Society Press, 1986: 67-83.
9. Kent DL, Shortliffe EH, Carlson RW, et al. Improvements in data collection through physician use of a computer-based chemotherapy treatment consultant. J Clin Oncol 1985; 3: 1409-17.
10. Lussier YA, Maksud M, Desruisseaux B, et al. PureMD: a computerized patient record software for direct data entry by physicians using a keyboard-free pen-based portable computer. In: Frisse ME, ed. *Sixteenth SCAMC*. New York: McGraw-Hill 1992; 261-4.
11. Kuhn K, Gaus W, Wechsler JG, et al. Structured reporting of medical findings: Evaluation of a system in gastroenterology. Meth Inform Med 1992; 31: 268-74.
12. Bell DS, and Greenes RA. Evaluation of UltraSTAR: Performance of a collaborative structured data entry system. In: Ozbolt JG, ed. *Eighteenth SCAMC*. Philadelphia: Hanley&Belfus, 1994: 216-22.
13. Gouveia-Oliveira A, Raposo VD, Salgado NC, et al. Longitudinal comparative study on the influence of computers on reporting of clinical data. Endoscopy 1991; 23: 334-37.
14. Van Ginneken AM, Stam H, and Moorman PW. A multi-strategy approach for medical records of specialists. In: Greenes RA et al, eds. *Medinfo 95*. Edmonton: HC & CC, 1995: 290-4.
15. Moorman PW, Van Ginneken AM, Van der Lei J, and van Bemmel JH. A model for structured data entry based on explicit descriptional knowledge. Meth Inform Med 1994; 33: 454-63.
16. Moorman PW, Van Ginneken AM, Van der Lei J, et al. The contents of free-text endoscopy reports: An inventory and evaluation by peers. Endoscopy 1994; 26: 531-8.
17. Moorman PW, Siersema PD, and Van Ginneken AM. Descriptive features of gastric ulcers: Do endoscopists agree on what they see?

Gastrointest Endosc 1995; 42: (in press).

18. Rind DM, and Safran C. Real and imagined barriers to an electronic medical record. In: Safran C, ed. *Seventeenth SCAMC*. New York: McGraw-Hill, 1993: 74-8
19. Van der Lei J, Duisterhout JS, Westerhof HP, et al. The introduction of computer-based patient records in The Netherlands. Ann Intern Med 1993; 119: 1036-41.

Appendix

Contents of the gastroscopy descriptional knowledgebase

In this appendix, the content of the (Dutch) gastroscopy descriptional knowledgebase is listed.

All 415 **concepts** are listed in **bold type-face**. The concepts are listed in alphabetical order, except for 'Gastroscopy', which is the top-concept. After each concept, its properties are listed between brackets. When a concept can exist multiple times, its plural form is also listed.

After each concept, first all relations with its children are listed in normal type-face.

Thereafter, in *italic type-face* all relations with the concept's parents are listed.

Relations and properties are listed in abbreviations. The full terms are as follows:

Relations

<HF> = Has Feature
<HS> = Has Specialisation
<HV> = Has Value
<EC> = Exclude Choice
<PC> = Preset Choice
<RT> = Refers To

Properties

absen = absent
multi = multiple
combi = combination
norma = normal

Gastroscopie []
<HF> algemeen
<HF> verslag
<HF> verrichtte handelingen
<HF> conclusie

accesoire papil [absen]
<HF> grootte
geidentificeerd <HF>
achterwand []
maagzijde <HS>
deel bulbus <HS>
afgenomen []
lumen diameter <HF>
licht reflex <HF>
elasticiteit van de wand <HF>
villi tekening <HF>
vaattekening <HF>
afgrensbaarheid tov omgeving []
<HS> scherp afgrensbaar van de omgeving
<HS> vaag afgrensbaar van de omgeving
erosie <HF>
ulcus <HF>
nodulaire afwijking <HF>
poliep <HF>
tumor <HF>

afspoelbaar []
afspoelbaarheid <HS>
afspoelbaarheid []
<HS> afspoelbaar
<HS> niet afspoelbaar
gebied met beslag <HF>
afstand []
<HF> van bovenrand tot tandenrij
<HF> tot bovenste oesofagus sfincter
<HF> tot onderste oesofagus sfincter
<HF> tot angulus
<HF> tot pylorus
<HF> tot papil van vater
slokdarm lokatie <HS>
maag lokatie <HS>
duodenum lokatie <HS>
slijmvlies <EC>
afstand tot tandenrij []
<HV> cm
scoop opgevoerd tot <HF>
hiatus <HF>
hiatus valt samen met de slokdarm-maag slijmvlies overgang <HF>
naar barrett epitheel <HF>
naar cylindrisch epitheel van maag <HF>

hiatus valt samen met de slokdarm-maag slijmvlies overgang <HF>
overgang barrett epitheel naar cylindrisch epitheel <HF>
afwezig []
peristaltiek <HF>
relaxatie <HF>
villi tekening <HF>
algemeen []
Gastroscopie <HF>
alle varices []
% van het aantal varices <HS>
anatomische duodenum lokatie []
<HS> in
<HS> tot en met
duodenum lokatie <HS>
scoop opgevoerd tot <PC>
anatomische maag lokatie []
<HS> gehele maag
<HS> cardia
<HS> fundus
<HS> corpus
<HS> angulus
<HS> antrum
<HS> prepylorisch antrum
<HS> pylorus
maag lokatie <HS>
scoop opgevoerd tot <PC>
anatomische slokdarm lokatie []
<HS> gehele slokdarm
<HS> bovenste derde deel van de slokdarm
<HS> middelste derde deel van de slokdarm
<HS> onderste derde deel van de slokdarm
<HS> omgeving z-lijn
slokdarm lokatie <HS>
scoop opgevoerd tot <PC>
andere afwijking []
<HS> vesikel
<HS> bulla
<HF> lokatie
<HF> grootte
<HF> kleur vocht
<HF> groepering
<HF> confluentie
bevindingen slokdarm <HF>
bevindingen duodenum <HF>
bevindingen maag <HF>
angiodysplasie [multi (angiodysplasien), absen]
<HF> grootte
<HF> bloedingsstigmata
vasculaire afwijking <HS>
angulus []
anatomische maag lokatie <HS>
antrum []
anatomische maag lokatie <HS>
asymmetrisch []
positie <HF>
asymmetrisch vervormd []
vorm pylorus <HS>
av-malformatie [multi (av-malformaties), absen]
<HF> grootte
<HF> bloedingsstigmata
vasculaire afwijking <HS>
beschrijving pylorus []
<HF> passeerbaarheid
<HF> vorm pylorus
<HF> positie
bevindingen maag <HF>
bevindingen duodenum <RT>
beslag [absen]
<HF> mate waarin bodem bedekt met beslag
<HF> kleur beslag
<HF> type beslag
bodem <HF>
bevindingen duodenum [norma]
<HF> inhoud
<HF> lumen diameter
<HF> plooi
<HF> stenose
<HF> slijmvlies
<HF> gebied met beslag
<HF> plekjes
<HF> defect
<HF> verheven afwijking
<HF> vasculaire afwijking
<HF> andere afwijking
<HF> papil van vater
<RT> beschrijving pylorus
<PC> duodenum lokatie
<PC> grootte tov het lumen
<EC> ring van schatzki
<EC> glycogeen stapelings plekje
<EC> plooi verstrijking
<EC> granulair patroon
verslag <HF>
bevindingen maag [norma]
<HF> inhoud
<HF> lumen diameter
<HF> elasticiteit van de wand
<HF> vorm maag
<HF> peristaltiek
<HF> slijmvlies
<HF> plooi
<HF> gebied met beslag
<HF> plekjes
<HF> defect
<HF> verheven afwijking
<HF> vasculaire afwijking
<HF> andere afwijking
<HF> beschrijving pylorus
<RT> hiatus
<PC> maag lokatie
<PC> varix grootte in maag
<EC> grootte tov de circumferentie
<EC> glycogeen stapelings plekje
<EC> villi tekening

<EC> sludge
<EC> helder vocht
<EC> endoprothese
verslag <HF>
bevindingen slokdarm [norma]
<HF> hiatus en z-lijn
<HF> peristaltiek
<HF> sfincters
<HF> inhoud
<HF> lumen diameter
<HF> stenose
<HF> slijmvlies
<HF> gebied met beslag
<HF> plekjes
<HF> defect
<HF> verheven afwijking
<HF> vasculaire afwijking
<HF> andere afwijking
<PC> slokdarm lokatie
<PC> grootte tov het lumen
<EC> tot papil van vater
<EC> tot pylorus
<EC> tot angulus
<EC> villi tekening
<EC> sludge
<EC> helder vocht
<EC> endoprothese
verslag <HF>
bezoar [absen]
vaste inhoud <HS>
secretie uit papil <EC>
binnenkant []
<HF> regelmatigheid
<HF> inhoud
<EC> lokatie
<EC> endoprothese
divertikel <HF>
blauw []
kleur <HF>
blauwer dan de omgeving []
kleur tov de omgeving <HS>
bleekheid []
<HF> type
<EC> lineair
kleur slijmvlies <HS>
bleker dan de omgeving []
kleur tov de omgeving <HS>
bloed [absen]
<HF> versheid
vloeibare inhoud <HS>
bloedingspatroon []
<HS> sijpelend
<HS> stromend
<HS> spuitend
<HS> massaal
spontaan bloedend <HF>
stenose <PC>
slijmvlies <PC>
fistel <PC>
divertikel <PC>
vasculaire ectasie <PC>
papil van vater <PC>
bloedingsplaats []
<HS> uit bodem
<HS> uit rand
spontaan bloedend <HF>
fissuur <EC>
perforatie <EC>
traumatische lesie <EC>
verheven afwijking <EC>
vasculaire afwijking <EC>
omgevend mucosa <EC>
rood plekje (niet nader te duiden) <EC>
bloedingsstigmata []
<HS> niet bloedend
<HS> makkelijk bloedend bij manipulatie
<HS> spontaan bloedend
<HS> tekenen van voorgaande bloeding
web <HF>
ring van schatzki <HF>
slijmvlies <HF>
(fibrotische) strictuur <HF>
naad strictuur <HF>
omgevend mucosa <HF>
ulcus <HF>
erosie <HF>
fissuur <HF>
perforatie <HF>
traumatische lesie <HF>
nodulaire afwijking <HF>
poliep <HF>
tumor <HF>
varix <HF>
hematoom <HF>
hemangioom <HF>
angiodysplasie <HF>
tele-angiectasie <HF>
av-malformatie <HF>
fistel <HF>
divertikel <HF>
vasculaire ectasie <HF>
extrensieke compressie <HF>
geidentificeerd <HF>
rood plekje (niet nader te duiden) <HF>
bodem []
<HF> beslag
<HF> regelmatigheid
<HF> granulatie weefsel
ulcus <HF>
erosie <HF>
bodem van bulbus []
deel bulbus <HS>
boven z-lijn []
omgeving z-lijn <HS>
bovenste derde deel van de slokdarm []
anatomische slokdarm lokatie <HS>

bovenste sfincter []
sfincters <HS>
bruin []
kleur beslag <HS>
kleur <HF>
bruin - zwart - melanosis []
kleur slijmvlies <HS>
bulbus []
<HF> deel bulbus
<PC> deel bulbus
in <HF>
tot en met <HF>
bulla [multi (bullae), absen]
andere afwijking <HS>
cardia []
anatomische maag lokatie <HS>
cascade maag []
vorm maag <HS>
catheter []
corpus alienum <HF>
circulair []
grootte tov de circumferentie <HS>
vorm <HS>
cm []
afstand tot tandenrij <HV>
grootte papillotomie opening <HV>
lengte <HV>
grootte <HV>
van bovenrand tot tandenrij <HV>
tot bovenste oesofagus sfincter <HV>
tot onderste oesofagus sfincter <HV>
tot angulus <HV>
tot pylorus <HV>
tot papil van vater <HV>
diepte <HV>
diameter van de ingang <HV>
diameter <HV>
slijmvlies overgang boven de hiatus <HV>
cobble stone patroon []
slijmvlies patroon <HS>
conclusie []
Gastroscopie <HF>
confluent []
confluentie <HS>
confluentie []
<HS> confluent
<HS> nonconfluent
gebied met beslag <HF>
andere afwijking <HF>
erosie <HF>
consistentie []
<HS> elastisch
<HS> vast
web <HF>
ring van schatzki <HF>
(fibrotische) strictuur <HF>
naad strictuur <HF>
tumor <HF>
extrensieke compressie <HF>
corpus []
<HF> deel
<PC> deel
anatomische maag lokatie <HS>
corpus alienum [multi (corpora aliena), absen]
<HF> munt
<HF> noot
<HF> drain
<HF> catheter
vaste inhoud <HS>
dak van bulbus []
deel bulbus <HS>
de tegenoverliggende darmwand eroderend []
endoprothese <HF>
deel []
<HS> proximale deel
<HS> middelste deel
<HS> distale deel
corpus <HF>
duodenum horizontale <HF>
corpus <PC>
duodenum horizontale <PC>
deel bulbus []
<HS> voorwand
<HS> achterwand
<HS> dak van bulbus
<HS> bodem van bulbus
bulbus <HF>
bulbus <PC>
scoop opgevoerd tot <EC>
defect []
<HS> ulcus
<HS> erosie
<HS> fissuur
<HS> divertikel
<HS> fistel
<HS> perforatie
<HS> litteken
<HS> traumatische lesie
<HF> lokatie
<EC> gehele slokdarm
<EC> gehele maag
bevindingen slokdarm <HF>
bevindingen maag <HF>
bevindingen duodenum <HF>
delle [absen]
nodulaire afwijking <HF>
diameter []
<HF> grootte beschrijving
<HV> mm
<HV> cm
fistel <HF>
diameter van de ingang []
<HV> mm
<HV> cm
divertikel <HF>
diep []

diepte beschrijving <HS>
diepte []
<HF> diepte beschrijving
<HV> mm
<HV> cm
ulcus <HF>
fissuur <HF>
divertikel <HF>
perforatie <HF>
traumatische lesie <HF>
diepte beschrijving []
<HS> ondiep
<HS> matig diep
<HS> diep
<HS> door de wand geperforeerd
diepte <HF>
diffuus []
type <HS>
diffuus verspreid []
groepering <HS>
dik []
dikte <HS>
dikte []
<HS> smal
<HS> dik
gesteeld <HF>
distaal van de papil van vater []
duodenum descendens <HF>
distale deel []
deel <HS>
divertikel [multi (divertikels), absen]
<HF> diepte
<HF> diameter van de ingang
<HF> binnenkant
<HF> bloedingsstigmata
<PC> bloedingspatroon
<EC> visible vessel
defect <HS>
door de wand geperforeerd []
diepte beschrijving <HS>
dorsaal []
slokdarmzijde <HS>
drain []
corpus alienum <HF>
drie kwart circulair []
grootte tov de circumferentie <HS>
"dubbele pylorus" []
vorm pylorus <HS>
duidelijk blauw []
varices kleur <HS>
duodenum descendens []
<HF> proximaal van de papil van vater
<HF> distaal van de papil van vater
<HF> tegenover de papil van vater
<HF> rond de papil van vater
in <HF>
tot en met <HF>
duodenum horizontale []
<HF> deel
<PC> deel
in <HF>
tot en met <HF>
duodenum lokatie []
<HS> anatomische duodenum lokatie
<HS> afstand
<EC> van bovenrand tot tandenrij
<EC> tot bovenste oesofagus sfincter
<EC> tot onderste oesofagus sfincter
<EC> tot angulus
<EC> tot pylorus
lokatie <HS>
scoop opgevoerd tot <HF>
bevindingen duodenum <PC>
ecchymose [multi (ecchymosen), absen]
plekjes <HS>
een achtste circulair []
grootte tov de circumferentie <HS>
een derde circulair []
grootte tov de circumferentie <HS>
eenvoudig te passeren met scoop []
passeerbaarheid <HS>
elasticiteit van de wand [combi]
<HF> normaal
<HF> afgenomen
<HF> lokatie
bevindingen maag <HF>
elastisch []
consistentie <HS>
endoprothese [absen]
<HF> uitvloed uit endoprothese
<HF> los in lumen liggend
<HF> de tegenoverliggende darmwand eroderend
geidentificeerd <HF>
soorten inhoud <RT>
bevindingen maag <EC>
bevindingen slokdarm <EC>
binnenkant <EC>
ernst []
<HS> mild
<HS> matig
<HS> ernstig
roodheid <HF>
gezwollen <HF>
ernstig []
ernst <HS>
erosie [multi (erosies), absen]
<HF> verhevenheid
<HF> grootte
<HF> vorm
<HF> bodem
<HF> bloedingsstigmata
<HF> afgrensbaarheid tov omgeving
<HF> omgevend mucosa
<HF> confluentie
<EC> hematine beslag
defect <HS>

extrensieke compressie [multi (extrensieke compressies), absen]
<HF> grootte
<HF> passeerbaarheid
<HF> oppervlak
<HF> kleur tov de omgeving
<HF> consistentie
<HF> bloedingsstigmata
<HF> pulsatie
verheven afwijking <HS>
stenose <RT>

faeces [absen]
<HF> hoeveelheid
soorten inhoud <HS>

fibrine [absen]
type beslag <HS>

(fibrotische) strictuur [multi (stricturen), absen]
<HF> lokatie
<HF> lengte
<HF> grootte
<HF> passeerbaarheid
<HF> oppervlak
<HF> consistentie
<HF> kleur tov de omgeving
<HF> bloedingsstigmata
<PC> grootte tov de circumferentie
stenose <HS>

fissuur [multi (fissuren), absen]
<HF> grootte
<HF> vorm
<HF> diepte
<HF> bloedingsstigmata
<HF> omgevend mucosa
<EC> visible vessel
<EC> bloedingsplaats
defect <HS>

fistel [multi (fistels), absen]
<HF> diameter
<HF> richting
<HF> omgevend mucosa
<HF> bloedingsstigmata
<PC> bloedingspatroon
<EC> visible vessel
defect <HS>

fundus []
anatomische maag lokatie <HS>

gal [absen]
vloeibare inhoud <HS>

gebied met beslag [multi (gebieden met beslag), absen]
<HF> lokatie
<HF> grootte
<HF> kleur
<HF> confluentie
<HF> afspoelbaarheid
bevindingen slokdarm <HF>
bevindingen maag <HF>
bevindingen duodenum <HF>

gedeelte []
<HS> gedeeltelijk
<HS> geheel
ondermijnd <HF>
opgehoogd <HF>
verstrijkend bij insufflatie <HF>
onregelmatig <HF>

gedeeltelijk []
gedeelte <HS>

gedeeltelijk bedekt met beslag []
mate waarin bodem bedekt met beslag <HS>

gedilateerd []
lumen diameter <HF>

geel []
kleur beslag <HS>
kleur <HF>

gefixeerd []
mobiliteit <HS>

geheel []
gedeelte <HS>

geheel bedekt met beslag []
mate waarin bodem bedekt met beslag <HS>

gehele dunne darm []
scoop opgevoerd tot <EC>

gehele maag []
anatomische maag lokatie <HS>
scoop opgevoerd tot <EC>
defect <EC>
verheven afwijking <EC>
stenose <EC>

gehele slokdarm []
anatomische slokdarm lokatie <HS>
scoop opgevoerd tot <EC>
stenose <EC>
defect <EC>
verheven afwijking <EC>

geïdentificeerd []
<HF> grootte
<HF> oppervlak
<HF> secretie uit papil
<HF> bloedingsstigmata
<HF> kleur slijmvlies
<HF> status na papillotomie
<HF> endoprothese
<HF> accesoire papil
papil van vater <HF>

geïndureerd []
induratie <HS>

gelobuleerd []
oppervlak <HF>

gesteeld []
<HF> dikte
<HF> steel lengte
steel <HS>

gezwollen []
<HF> ernst
<HF> type
zwelling <HS>

glad []
oppervlak <HF>
glycogeen stapelings plekje [multi (glycogeen stapelings plekjes), absen]
plekjes <HS>
bevindingen maag <EC>
bevindingen duodenum <EC>
granulair patroon []
slijmvlies patroon <HS>
bevindingen duodenum <EC>
granulatie weefsel [absen]
bodem <HF>
grijs []
kleur beslag <HS>
kleur <HF>
groen []
kleur <HF>
groepering []
<HS> diffuus verspreid
<HS> in groepjes
andere afwijking <HF>
groot []
grootte beschrijving <HS>
grootte []
<HF> grootte beschrijving
<HF> grootte tov de circumferentie
<HV> mm
<HV> cm
varix grootte in maag <HS>
stolsel <HF>
web <HF>
(fibrotische) strictuur <HF>
naad strictuur <HF>
ulcus <HF>
erosie <HF>
fissuur <HF>
perforatie <HF>
litteken <HF>
traumatische lesie <HF>
nodulaire afwijking <HF>
poliep <HF>
tumor <HF>
vasculaire ectasie <HF>
hematoom <HF>
hemangioom <HF>
angiodysplasie <HF>
tele-angiectasie <HF>
av-malformatie <HF>
andere afwijking <HF>
extrensieke compressie <HF>
ring van schatzki <HF>
gebied met beslag <HF>
geidentificeerd <HF>
accesoire papil <HF>
paraoesofageale hernia <HF>
grootte beschrijving []
<HS> puntgroot
<HS> klein
<HS> matig groot
<HS> groot
<HS> zeer groot
grootte <HF>
diameter <HF>
grootte papillotomie opening <HF>
stolsel <PC>
papil van vater <PC>
paraoesofageale hernia <PC>
grootte papillotomie opening []
<HF> grootte beschrijving
<HV> mm
<HV> cm
status na papillotomie <HF>
grootte tov de circumferentie []
<HS> een achtste circulair
<HS> kwart circulair
<HS> een derde circulair
<HS> half circulair
<HS> drie kwart circulair
<HS> circulair
grootte <HF>
web <PC>
(fibrotische) strictuur <PC>
naad strictuur <PC>
ring van schatzki <PC>
bevindingen maag <EC>
grootte tov het lumen []
<HS> klein, nauwelijks zichtbaar (graad 1)
<HS> 1/4 van het lumen innemend (graad 2)
<HS> 1/2 van het lumen innemend (graad 3)
<HS> meer dan 1/2 van het lumen innemend (graad 4)
varix grootte <HS>
bevindingen slokdarm <PC>
bevindingen duodenum <PC>
grootte tov normaal []
<HS> normale grootte
<HS> kleiner dan normaal
<HS> groter dan normaal
plooi <HF>
grote curvatuur []
maagzijde <HS>
grote hoeveelheid []
hoeveelheid <HS>
groter dan normaal []
grootte tov normaal <HS>
half circulair []
grootte tov de circumferentie <HS>
hematine plekje (black spot) [multi (hematine plekjes), absen]
plekjes <HS>
hemorhagisch plekje [multi (hemorhagische plekjes), absen]
plekjes <HS>
hechting [multi (hechtingen), absen]
vaste inhoud <HS>
helder vocht [absen]

vloeibare inhoud <HS>
bevindingen maag <EC>
bevindingen slokdarm <EC>
hemangioom [multi (hemangiomen), absen]
<HF> grootte
<HF> bloedingsstigmata
vasculaire afwijking <HS>
verheven afwijking <RT>
hematine [absen]
type beslag <HS>
hematine beslag [absen]
tekenen van voorgaande bloeding <HF>
ulcus <EC>
erosie <EC>
hematoom [multi (hematomen), absen]
<HF> grootte
<HF> bloedingsstigmata
vasculaire afwijking <HS>
hiatus []
<HF> sluit rondom de scoop
<HF> laat ruimte vrij rond de scoop
<HF> afstand tot tandenrij
hiatus en z-lijn <HS>
bevindingen maag <RT>
hiatus en z-lijn []
<HS> hiatus
<HS> overgang plaveisel epitheel
<HS> hiatus valt samen met de slokdarm-maag slijmvlies overgang
<HS> slijmvlies overgang boven de hiatus
<HS> paraoesofageale hernia
bevindingen slokdarm <HF>
hiatus valt samen met de slokdarm-maag slijmvlies overgang []
<HF> afstand tot tandenrij
hiatus en z-lijn <HS>
hoeveelheid []
<HS> minimale hoeveelheid
<HS> middelmatige hoeveelheid
<HS> grote hoeveelheid
voedsel <HF>
faeces <HF>
parasieten <HF>
vloeibare inhoud <HF>
hoogte []
<HS> niet opgehoogd
<HS> opgehoogd
rand <HF>
houden op voor het bereiken van de rand []
plooi verloop <HF>
hour-glass (bilocular) maag []
vorm maag <HS>
in []
<HF> bulbus
<HF> duodenum descendens
<HF> duodenum horizontale
<HF> nabij het ligament van treitz
<HF> proximale deel jejunum
anatomische duodenum lokatie <HS>
scoop opgevoerd tot <PC>
inhoud <PC>
in groepjes []
groepering <HS>
induratie []
<HS> geindureerd
<HS> niet geindureerd
rand <HF>
omgevend mucosa <HF>
inhoud [absen]
<HF> soorten inhoud
<PC> soorten inhoud
<PC> in
bevindingen slokdarm <HF>
binnenkant <HF>
bevindingen maag <HF>
bevindingen duodenum <HF>
intakt []
oppervlak <HF>
introductie []
<HF> vlotheid
<PC> vlotheid
introductie en opvoering <HS>
introductie en opvoering []
<HS> introductie
<HS> scoop opgevoerd tot
<HS> scoop opvoering
verslag <HF>
klein []
grootte beschrijving <HS>
klein, nauwelijks zichtbaar (graad 1) []
grootte tov het lumen <HS>
kleine curvatuur []
maagzijde <HS>
kleiner dan normaal []
grootte tov normaal <HS>
kleur []
<HF> wit
<HF> grijs
<HF> bruin
<HF> rood
<HF> zwart
<HF> blauw
<HF> geel
<HF> groen
kleur beschrijving <HS>
gebied met beslag <HF>
kleur vocht <HF>
kleur vocht <PC>
kleur beschrijving []
<HS> kleur
<HS> kleur tov de omgeving
poliep <HF>
tumor <HF>
nodulaire afwijking <HF>
kleur beslag []
<HS> wit

<HS> geel
<HS> grijs
<HS> bruin
<HS> rood
<HS> zwart
beslag <HF>
kleur gelijk aan de omgeving []
kleur tov de omgeving <HS>
varices kleur <HS>
kleur slijmvlies []
<HS> normale kleur
<HS> roodheid
<HS> bleekheid
<HS> bruin - zwart - melanosis
slijmvlies <HF>
geidentificeerd <HF>
kleur tov de omgeving []
<HS> kleur gelijk aan de omgeving
<HS> bleker dan de omgeving
<HS> roder dan de omgeving
<HS> blauwer dan de omgeving
kleur beschrijving <HS>
web <HF>
ring van schatzki <HF>
(fibrotische) strictuur <HF>
naad strictuur <HF>
litteken <HF>
extrensieke compressie <HF>
kleur vocht []
<HF> kleur
<PC> kleur
andere afwijking <HF>
kort []
steel lengte <HS>
kwart circulair []
grootte tov de circumferentie <HS>
laat ruimte vrij rond de scoop []
hiatus <HF>
lang []
steel lengte <HS>
lengte []
<HV> mm
<HV> cm
web <HF>
ring van schatzki <HF>
(fibrotische) strictuur <HF>
naad strictuur <HF>
licht blauw []
varices kleur <HS>
licht reflex []
<HF> normaal
<HF> afgenomen
slijmvlies <HF>
lineair []
type <HS>
vorm <HS>
bleekheid <EC>
links lateraal []
slokdarmzijde <HS>
litteken [multi (littekens), absen]
<HF> grootte
<HF> vorm
<HF> verhevenheid
<HF> kleur tov de omgeving
defect <HS>
lokatie []
<HS> slokdarm lokatie
<HS> maag lokatie
<HS> duodenum lokatie
soorten inhoud <HF>
web <HF>
ring van schatzki <HF>
gebied met beslag <HF>
slijmvlies <HF>
(fibrotische) strictuur <HF>
naad strictuur <HF>
plekjes <HF>
defect <HF>
verheven afwijking <HF>
vasculaire afwijking <HF>
andere afwijking <HF>
elasticiteit van de wand <HF>
lumen diameter <HF>
plooi <HF>
binnenkant <EC>
papil van vater <EC>
lopen door tot in de krater []
plooi verloop <HF>
los in lumen liggend []
endoprothese <HF>
lumen diameter [combi]
<HF> afgenomen
<HF> normaal
<HF> gedilateerd
<HF> lokatie
bevindingen slokdarm <HF>
bevindingen maag <HF>
bevindingen duodenum <HF>
vorm maag <RT>
maag lokatie []
<HS> anatomische maag lokatie
<HS> zijde
<HS> afstand
<PC> maagzijde
<EC> tot bovenste oesofagus sfincter
<EC> tot papil van vater
lokatie <HS>
scoop opgevoerd tot <HF>
bevindingen maag <PC>
maag sekreet [absen]
vloeibare inhoud <HS>
secretie uit papil <EC>
uitvloed uit endoprothese <EC>
maagzijde []
<HS> voorwand
<HS> achterwand

<HS> grote curvatuur
<HS> kleine curvatuur
zijde <HS>
maag lokatie <PC>
makkelijk bloedend bij manipulatie []
bloedingsstigmata <HS>
massaal []
bloedingspatroon <HS>
mate waarin bodem bedekt met beslag []
<HS> niet bedekt met beslag
<HS> gedeeltelijk bedekt met beslag
<HS> geheel bedekt met beslag
beslag <HF>
matig []
ernst <HS>
matig diep []
diepte beschrijving <HS>
matig groot []
grootte beschrijving <HS>
meer dan 1/2 van het lumen innemend (graad 4) []
grootte tov het lumen <HS>
met barrett uitlopers in plaveisel epitheel []
naar barrett epitheel <HF>
met enige moeite []
vlotheid <HS>
met enige pijn gepaard gaand []
vlotheid <HS>
met maagslijmvlies uitlopers []
naar cylindrisch epitheel van maag <HF>
met reflux []
peristaltiek <HF>
met veel pijn gepaard gaand []
vlotheid <HS>
middelmatige hoeveelheid []
hoeveelheid <HS>
middelste deel []
deel <HS>
middelste derde deel van de slokdarm []
anatomische slokdarm lokatie <HS>
mild []
ernst <HS>
minimale hoeveelheid []
hoeveelheid <HS>
mm []
lengte <HV>
grootte <HV>
diepte <HV>
diameter van de ingang <HV>
diameter <HV>
grootte papillotomie opening <HV>
slijmvlies overgang boven de hiatus <HV>
mobiel []
mobiliteit <HS>
mobiliteit []
<HS> mobiel
<HS> gefixeerd
poliep <HF>
tumor <HF>
munt []
corpus alienum <HF>
papil van vater <EC>
na enig aandringen te passeren met scoop []
passeerbaarheid <HS>
naad strictuur [multi (naad stricturen), absen]
<HF> lokatie
<HF> lengte
<HF> grootte
<HF> passeerbaarheid
<HF> oppervlak
<HF> consistentie
<HF> kleur tov de omgeving
<HF> bloedingsstigmata
<PC> grootte tov de circumferentie
stenose <HS>
naar barrett epitheel []
<HF> afstand tot tandenrij
<HF> scherp
<HF> met barrett uitlopers in plaveisel epitheel
<HF> overgang barrett epitheel naar cylindrisch epitheel
overgang plaveisel epitheel <HF>
naar cylindrisch epitheel van maag []
<HF> afstand tot tandenrij
<HF> scherp
<HF> met maagslijmvlies uitlopers
overgang plaveisel epitheel <HF>
nabij het ligament van treitz []
in <HF>
tot en met <HF>
necrose [absen]
type beslag <HS>
necrotisch []
oppervlak <HF>
niet afspoelbaar []
afspoelbaarheid <HS>
niet bedekt met beslag []
mate waarin bodem bedekt met beslag <HS>
niet bloedend []
bloedingsstigmata <HS>
niet geidentificeerd []
papil van vater <HF>
niet geindureerd []
induratie <HS>
niet gezwollen []
zwelling <HS>
niet ondermijnd []
ondermijning <HS>
niet opgehoogd []
hoogte <HS>
niet pulserend []
pulsatie <HS>
niet te passeren met scoop, wel met voerdraad []
passeerbaarheid <HS>
niet te passeren met voerdraad []
passeerbaarheid <HS>

niet verheven []
verhevenheid <HS>
niet verstrijkend bij insufflatie []
plooi verstrijking <HS>
wegblaasbaarheid <HS>
nodulair patroon []
slijmvlies patroon <HS>
nodulaire afwijking [multi (nodulaire afwijkingen), absen]
<HF> grootte
<HF> oppervlak
<HF> kleur beschrijving
<HF> delle
<HF> bloedingsstigmata
<HF> afgrensbaarheid tov omgeving
<HF> omgevend mucosa
verheven afwijking <HS>
nonconfluent []
confluentie <HS>
noot []
corpus alienum <HF>
papil van vater <EC>
normaal []
peristaltiek <HF>
relaxatie <HF>
lumen diameter <HF>
licht reflex <HF>
elasticiteit van de wand <HF>
villi tekening <HF>
vaattekening <HF>
normaal patroon []
slijmvlies patroon <HS>
normale grootte []
grootte tov normaal <HS>
normale kleur []
kleur slijmvlies <HS>
omgevend mucosa <HF>
normale vorm []
vorm maag <HS>
omgevend mucosa []
<HF> normale kleur
<HF> roodheid
<HF> zwelling
<HF> induratie
<HF> verhevenheden
<HF> bloedingsstigmata
<EC> visible vessel
<EC> bloedingsplaats
ulcus <HF>
erosie <HF>
fistel <HF>
fissuur <HF>
nodulaire afwijking <HF>
poliep <HF>
tumor <HF>
omgeving z-lijn []
<HS> boven z-lijn
<HS> op z-lijn
<HS> onder z-lijn
anatomische slokdarm lokatie <HS>
ring van schatzki <PC>
onder z-lijn []
omgeving z-lijn <HS>
ondermijnd []
<HF> gedeelte
ondermijning <HS>
ondermijning []
<HS> niet ondermijnd
<HS> ondermijnd
rand <HF>
onderste derde deel van de slokdarm []
anatomische slokdarm lokatie <HS>
onderste sfincter []
sfincters <HS>
ondiep []
diepte beschrijving <HS>
onregelmatig []
<HF> gedeelte
regelmatigheid <HS>
oppervlak <HF>
onregelmatige vorm []
vorm <HS>
op z-lijn []
omgeving z-lijn <HS>
opgehoogd []
<HF> gedeelte
hoogte <HS>
oppervlak []
<HF> regelmatig
<HF> glad
<HF> gelobuleerd
<HF> intakt
<HF> onregelmatig
<HF> ulceratief
<HF> necrotisch
<HF> villeus
web <HF>
ring van schatzki <HF>
(fibrotische) strictuur <HF>
naad strictuur <HF>
nodulaire afwijking <HF>
poliep <HF>
tumor <HF>
extrensieke compressie <HF>
geidentificeerd <HF>
oud []
versheid <HS>
ovaal []
vorm <HS>
overgang barrett epitheel naar cylindrisch epitheel []
<HF> afstand tot tandenrij
naar barrett epitheel <HF>
overgang plaveisel epitheel []
<HF> naar cylindrisch epitheel van maag
<HF> naar barrett epitheel

hiatus en z-lijn <HS>
papil van vater []
<HF> geïdentificeerd
<HF> niet geïdentificeerd
<PC> grootte beschrijving
<PC> bloedingspatroon
<EC> visible vessel
<EC> lokatie
<EC> munt
<EC> noot
bevindingen duodenum <HF>
paraoesofageale hernia [absen]
<HF> grootte
<PC> grootte beschrijving
hiatus en z-lijn <HS>
parasieten [combi, absen]
<HF> hoeveelheid
soorten inhoud <HS>
secretie uit papil <HS>
passeerbaarheid []
<HS> eenvoudig te passeren met scoop
<HS> na enig aandringen te passeren met scoop
<HS> niet te passeren met scoop, wel met voerdraad
<HS> niet te passeren met voerdraad
sfincters <HF>
web <HF>
ring van schatzki <HF>
(fibrotische) strictuur <HF>
naad strictuur <HF>
poliep <HF>
tumor <HF>
extrensieke compressie <HF>
beschrijving pylorus <HF>
perforatie [multi (perforaties), absen]
<HF> grootte
<HF> vorm
<HF> diepte
<HF> bloedingsstigmata
<EC> visible vessel
<EC> bloedingsplaats
defect <HS>
peristaltiek []
<HF> normaal
<HF> afwezig
<HF> verminderd
<HF> verhoogd
<HF> retrograad
<HF> met reflux
<HF> spasme
bevindingen slokdarm <HF>
bevindingen maag <HF>
petechie [multi (petechieen), absen]
plekjes <HS>
plekjes []
<HS> petechie
<HS> hemorhagisch plekje
<HS> hematine plekje (black spot)
<HS> ecchymose
<HS> glycogeen stapelings plekje
<HS> rood plekje (niet nader te duiden)
<HF> lokatie
bevindingen slokdarm <HF>
bevindingen maag <HF>
bevindingen duodenum <HF>
plooi [multi (plooien)]
<HF> lokatie
<HF> grootte tov normaal
<HF> zwelling
<HF> plooi verstrijking
bevindingen maag <HF>
bevindingen duodenum <HF>
plooi verloop []
<HF> lopen door tot in de krater
<HF> houden op voor het bereiken van de rand
ulcus <HF>
plooi verstrijking []
<HS> niet verstrijkend bij insufflatie
<HS> verstrijkend bij insufflatie
plooi <HF>
bevindingen duodenum <EC>
poliep [multi (pollepen), absen]
<HF> grootte
<HF> omgevend mucosa
<HF> steel
<HF> mobiliteit
<HF> kleur beschrijving
<HF> oppervlak
<HF> bloedingsstigmata
<HF> afgrensbaarheid tov omgeving
<HF> passeerbaarheid
verheven afwijking <HS>
stenose <RT>
positie []
<HF> symmetrisch
<HF> asymmetrisch
beschrijving pylorus <HF>
prepylorisch antrum []
anatomische maag lokatie <HS>
proximaal van de papil van vater []
duodenum descendens <HF>
proximale deel []
deel <HS>
proximale deel jejunum []
in <HF>
tot en met <HF>
pulsatie []
<HS> niet pulserend
<HS> pulserend
extrensieke compressie <HF>
pulserend []
pulsatie <HS>
puntgroot []
grootte beschrijving <HS>
pus [absen]
vloeibare inhoud <HS>
pylorus []

anatomische maag lokatie <HS>
rand []
<HF> hoogte
<HF> regelmatigheid
<HF> ondermijning
<HF> induratie
ulcus <HF>
recht []
verloop <HF>
rechts lateraal []
slokdarmzijde <HS>
regelmatig []
regelmatigheid <HS>
oppervlak <HF>
regelmatigheid []
<HS> regelmatig
<HS> onregelmatig
bodem <HF>
rand <HF>
binnenkant <HF>
relaxatie []
<HF> normaal
<HF> afwezig
sfincters <HF>
retrograad []
peristaltiek <HF>
richting []
<HS> richting trachea
<HS> richting bronchus
<HS> richting aorta
<HS> richting colon
<HS> richting pancreas
<HS> richting choledochus
<HS> richting galblaas
<HS> richting andere
fistel <HF>
richting andere []
richting <HS>
richting aorta []
richting <HS>
richting bronchus []
richting <HS>
richting choledochus []
richting <HS>
richting colon []
richting <HS>
richting galblaas []
richting <HS>
richting pancreas []
richting <HS>
richting trachea []
richting <HS>
ring van schatzki [multi (ringen van schatzki), absen]
<HF> lokatie
<HF> lengte
<HF> grootte
<HF> passeerbaarheid
<HF> oppervlak
<HF> consistentie
<HF> kleur tov de omgeving
<HF> bloedingsstigmata
<PC> omgeving z-lijn
<PC> grootte tov de circumferentie
<EC> zijde
stenose <HS>
bevindingen duodenum <EC>
rode spots < 3 mm (cherry red spots) [absen]
rode spots en streepjes op varices <HS>
rode spots > 3 mm (hemacystic spots) [absen]
rode spots en streepjes op varices <HS>
rode spots en streepjes op varices []
<HS> rode streepjes (red wale markings)
<HS> rode spots < 3 mm (cherry red spots)
<HS> rode spots > 3 mm (hemacystic spots)
<HF> % van het aantal varices
varix <HF>
rode streepjes (red wale markings) [absen]
rode spots en streepjes op varices <HS>
roder dan de omgeving []
kleur tov de omgeving <HS>
rond []
vorm pylorus <HS>
rond de papil van vater []
duodenum descendens <HF>
rood []
kleur beslag <HS>
kleur <HF>
rood plekje (niet nader te duiden) [multi (rode plekjes (niet nader te duiden))]
<HF> bloedingsstigmata
<EC> bloedingsplaats
<EC> visible vessel
plekjes <HS>
roodheid []
<HF> ernst
<HF> type
kleur slijmvlies <HS>
omgevend mucosa <HF>
scherp []
naar barrett epitheel <HF>
naar cylindrisch epitheel van maag <HF>
scherp afgrensbaar van de omgeving []
afgrensbaarheid tov omgeving <HS>
scoop opgevoerd tot []
<HF> slokdarm lokatie
<HF> maag lokatie
<HF> duodenum lokatie
<HF> afstand tot tandenrij
<PC> anatomische slokdarm lokatie
<PC> anatomische maag lokatie
<PC> anatomische duodenum lokatie
<PC> in
<EC> gehele slokdarm
<EC> gehele maag
<EC> gehele dunne darm

<EC> deel bulbus
introductie en opvoering <HS>
scoop opvoering []
<HF> vlotheid
<PC> vlotheid
introductie en opvoering <HS>
secretie uit papil [absen]
<HS> voedsel
<HS> vloeibare inhoud
<HS> vaste inhoud
<HS> parasieten
<EC> maag sekreet
<EC> speeksel
<EC> bezoar
<EC> tablet
geidentificeerd <HF>
semisessiel []
steel <HS>
serpigineus []
vorm <HS>
sessiel []
steel <HS>
sfincters []
<HS> bovenste sfincter
<HS> onderste sfincter
<HF> passeerbaarheid
<HF> relaxatie
bevindingen slokdarm <HF>
sijpelend []
bloedingspatroon <HS>
slijmvlies [combi]
<HF> lokatie
<HF> kleur slijmvlies
<HF> zwelling
<HF> slijmvlies patroon
<HF> villi tekening
<HF> bloedingsstigmata
<HF> licht reflex
<HF> vaattekening
<PC> bloedingspatroon
<EC> tekenen van voorgaande bloeding
<EC> afstand
bevindingen slokdarm <HF>
bevindingen maag <HF>
bevindingen duodenum <HF>
slijmvlies overgang boven de hiatus []
<HV> mm
<HV> cm
hiatus en z-lijn <HS>
slijmvlies patroon []
<HS> normaal patroon
<HS> granulair patroon
<HS> nodulair patroon
<HS> cobble stone patroon
slijmvlies <HF>
slokdarm lokatie []
<HS> anatomische slokdarm lokatie
<HS> zijde
<HS> afstand
<PC> slokdarmzijde
<EC> tot angulus
<EC> tot pylorus
<EC> tot papil van vater
lokatie <HS>
scoop opgevoerd tot <HF>
bevindingen slokdarm <PC>
slokdarmzijde []
<HS> links lateraal
<HS> rechts lateraal
<HS> ventraal
<HS> dorsaal
zijde <HS>
slokdarm lokatie <PC>
sludge [absen]
vloeibare inhoud <HS>
bevindingen maag <EC>
bevindingen slokdarm <EC>
sluit rondom de scoop []
hiatus <HF>
smal []
dikte <HS>
soorten inhoud []
<HS> voedsel
<HS> faeces
<HS> vloeibare inhoud
<HS> vaste inhoud
<HS> parasieten
<HF> lokatie
<RT> endoprothese
inhoud <HF>
inhoud <PC>
spasme []
peristaltiek <HF>
spastisch []
vernauwd <HF>
speeksel [absen]
vloeibare inhoud <HS>
secretie uit papil <EC>
uitvloed uit endoprothese <EC>
spontaan bloedend []
<HF> bloedingspatroon
<HF> bloedingsplaats
bloedingsstigmata <HS>
spuitend []
bloedingspatroon <HS>
status na papillotomie []
<HF> grootte papillotomie opening
geidentificeerd <HF>
steel []
<HS> sessiel
<HS> semisessiel
<HS> gesteeld
poliep <HF>
steel lengte []
<HS> lang
<HS> kort

gesteeld <HF>
steen [multi (stenen), absen]
vaste inhoud <HS>
stenose []
<HS> web
<HS> ring van schatzki
<HS> (fibrotische) strictuur
<HS> naad strictuur
<RT> tumor
<RT> poliep
<RT> extrensieke compressie
<PC> bloedingspatroon
<EC> visible vessel
<EC> gehele slokdarm
<EC> gehele maag
bevindingen slokdarm <HF>
bevindingen duodenum <HF>
stenotisch []
vernauwd <HF>
stervormig []
vorm <HS>
stolsel [absen]
<HF> grootte
<PC> grootte beschrijving
tekenen van voorgaande bloeding <HF>
stromend []
bloedingspatroon <HS>
symmetrisch []
positie <HF>
tablet [multi (tabletten), absen]
vaste inhoud <HS>
secretie uit papil <EC>
tegenover de papil van vater []
duodenum descendens <HF>
tekenen van voorgaande bloeding [absen]
<HF> stolsel
<HF> hematine beslag
<HF> visible vessel
bloedingsstigmata <HS>
slijmvlies <EC>
tele-angiectasie [multi (tele-angiectasien), absen]
<HF> grootte
<HF> bloedingsstigmata
vasculaire afwijking <HS>
toegenomen []
villi tekening <HF>
vaattekening <HF>
tortueus []
verloop <HF>
tot angulus []
<HV> cm
afstand <HF>
bevindingen slokdarm <EC>
slokdarm lokatie <EC>
duodenum lokatie <EC>
tot bovenste oesofagus sfincter []
<HV> cm
afstand <HF>
maag lokatie <EC>
duodenum lokatie <EC>
tot en met []
<HF> bulbus
<HF> duodenum descendens
<HF> duodenum horizontale
<HF> nabij het ligament van treitz
<HF> proximale deel jejunum
anatomische duodenum lokatie <HS>
tot onderste oesofagus sfincter []
<HV> cm
afstand <HF>
duodenum lokatie <EC>
tot papil van vater []
<HV> cm
afstand <HF>
bevindingen slokdarm <EC>
maag lokatie <EC>
slokdarm lokatie <EC>
tot pylorus []
<HV> cm
afstand <HF>
bevindingen slokdarm <EC>
slokdarm lokatie <EC>
duodenum lokatie <EC>
traumatische lesie [multi (traumatische lesies), absen]
<HF> grootte
<HF> vorm
<HF> diepte
<HF> bloedingsstigmata
<EC> bloedingsplaats
<EC> visible vessel
defect <HS>
tumor [multi (tumoren), absen]
<HF> grootte
<HF> consistentie
<HF> vorm
<HF> kleur beschrijving
<HF> oppervlak
<HF> passeerbaarheid
<HF> omgevend mucosa
<HF> afgrensbaarheid tov omgeving
<HF> bloedingsstigmata
<HF> mobiliteit
verheven afwijking <HS>
stenose <RT>
type []
<HS> vlekkerig
<HS> lineair
<HS> diffuus
roodheid <HF>
bleekheid <HF>
gezwollen <HF>
type beslag []
<HS> fibrine
<HS> hematine
<HS> necrose

beslag <HF>
uit bodem []
bloedingsplaats <HS>
uit rand []
bloedingsplaats <HS>
uitvloed uit endoprothese [absen]
<HF> vloeibare inhoud
<PC> vloeibare inhoud
<EC> speeksel
<EC> maag sekreet
endoprothese <HF>
ulceratief []
oppervlak <HF>
ulcus [multi (ulcera), absen]
<HF> grootte
<HF> vorm
<HF> diepte
<HF> bodem
<HF> rand
<HF> bloedingsstigmata
<HF> afgrensbaarheid tov omgeving
<HF> omgevend mucosa
<HF> plooi verloop
<EC> hematine beslag
defect <HS>
vaag afgrensbaar van de omgeving []
afgrensbaarheid tov omgeving <HS>
vaattekening []
<HF> normaal
<HF> afgenomen
<HF> toegenomen
slijmvlies <HF>
van bovenrand tot tandenrij []
<HV> cm
afstand <HF>
duodenum lokatie <EC>
varices kleur []
<HS> kleur gelijk aan de omgeving
<HS> licht blauw
<HS> duidelijk blauw
<HF> % van het aantal varices
varix <HF>
varix [multi (varices), absen]
<HF> varix grootte
<HF> varices kleur
<HF> verloop
<HF> rode spots en streepjes op varices
<HF> bloedingsstigmata
<HF> wegblaasbaarheid
vasculaire afwijking <HS>
verheven afwijking <RT>
varix grootte []
<HS> grootte tov het lumen
<HS> varix grootte in maag
varix <HF>
varix grootte in maag []
<HS> grootte
varix grootte <HS>
bevindingen maag <PC>
vasculaire afwijking []
<HS> varix
<HS> vasculaire ectasie
<HS> hematoom
<HS> hemangioom
<HS> angiodysplasie
<HS> tele-angiectasie
<HS> av-malformatie
<HF> lokatie
<EC> visible vessel
<EC> bloedingsplaats
bevindingen slokdarm <HF>
bevindingen maag <HF>
bevindingen duodenum <HF>
vasculaire ectasie [multi (vasculaire ectasieen), absen]
<HF> grootte
<HF> bloedingsstigmata
<PC> bloedingspatroon
<EC> visible vessel
vasculaire afwijking <HS>
vast []
consistentie <HS>
vaste inhoud []
<HS> steen
<HS> corpus alienum
<HS> bezoar
<HS> tablet
<HS> hechting
soorten inhoud <HS>
secretie uit papil <HS>
ventraal []
slokdarmzijde <HS>
verheven []
verhevenheid <HS>
verheven afwijking []
<HS> nodulaire afwijking
<HS> poliep
<HS> tumor
<HS> extrensieke compressie
<HF> lokatie
<RT> varix
<RT> hemangioom
<EC> gehele slokdarm
<EC> gehele maag
<EC> visible vessel
<EC> bloedingsplaats
bevindingen slokdarm <HF>
bevindingen maag <HF>
bevindingen duodenum <HF>
verhevenheden [absen]
omgevend mucosa <HF>
verhevenheid []
<HS> verheven
<HS> niet verheven
erosie <HF>
litteken <HF>

verhoogd []
peristaltiek <HF>
verloop []
<HF> tortueus
<HF> recht
varix <HF>
verminderd []
peristaltiek <HF>
vernauwd []
<HF> spastisch
<HF> stenotisch
vorm pylorus <HS>
verrichtte handelingen []
Gastroscopie <HF>
vers []
versheid <HS>
versheid []
<HS> oud
<HS> vers
bloed <HF>
verslag []
<HF> introductie en opvoering
<HF> bevindingen slokdarm
<HF> bevindingen maag
<HF> bevindingen duodenum
Gastroscopie <HF>
verstrijkend bij insufflatie []
<HF> gedeelte
plooi verstrijking <HS>
wegblaasbaarheid <HS>
verwijd []
vorm pylorus <HS>
vesikel [multi (vesikels), absen]
andere afwijking <HS>
villeus []
oppervlak <HF>
villi tekening []
<HF> normaal
<HF> afwezig
<HF> afgenomen
<HF> toegenomen
slijmvlies <HF>
bevindingen maag <EC>
bevindingen slokdarm <EC>
visible vessel [absen]
tekenen van voorgaande bloeding <HF>
stenose <EC>
fissuur <EC>
perforatie <EC>
traumatische lesie <EC>
verheven afwijking <EC>
vasculaire afwijking <EC>
omgevend mucosa <EC>
fistel <EC>
divertikel <EC>
vasculaire ectasie <EC>
papil van vater <EC>
rood plekje (niet nader te duiden) <EC>
vlekkerig []
type <HS>
vloeibare inhoud []
<HS> speeksel
<HS> maag sekreet
<HS> gal
<HS> bloed
<HS> pus
<HS> sludge
<HS> helder vocht
<HF> hoeveelheid
soorten inhoud <HS>
secretie uit papil <HS>
uitvloed uit endoprothese <HF>
uitvloed uit endoprothese <PC>
vlot []
vlotheid <HS>
vlotheid []
<HS> vlot
<HS> met enige moeite
<HS> met enige pijn gepaard gaand
<HS> met veel pijn gepaard gaand
introductie <HF>
scoop opvoering <HF>
introductie <PC>
scoop opvoering <PC>
voedsel [absen]
<HF> hoeveelheid
soorten inhoud <HS>
secretie uit papil <HS>
voorwand []
maagzijde <HS>
deel bulbus <HS>
vorm []
<HS> circulair
<HS> ovaal
<HS> lineair
<HS> serpigineus
<HS> stervormig
<HS> onregelmatige vorm
ulcus <HF>
erosie <HF>
fissuur <HF>
perforatie <HF>
litteken <HF>
traumatische lesie <HF>
tumor <HF>
vorm maag []
<HS> normale vorm
<HS> cascade maag
<HS> hour-glass (bilocular) maag
<RT> lumen diameter
bevindingen maag <HF>
vorm pylorus []
<HS> rond
<HS> asymmetrisch vervormd
<HS> verwijd
<HS> vernauwd

<HS> "dubbele pylorus"
beschrijving pylorus <HF>
web [multi (webben), absen]
<HF> lokatie
<HF> lengte
<HF> grootte
<HF> passeerbaarheid
<HF> oppervlak
<HF> consistentie
<HF> kleur tov de omgeving
<HF> bloedingsstigmata
<PC> grootte tov de circumferentie
stenose <HS>
wegblaasbaarheid []
<HS> niet verstrijkend bij insufflatie
<HS> verstrijkend bij insufflatie
varix <HF>
wit []
kleur beslag <HS>
kleur <HF>
zeer groot []
grootte beschrijving <HS>
zijde []
<HS> maagzijde
<HS> slokdarmzijde
slokdarm lokatie <HS>
maag lokatie <HS>
ring van schatzki <EC>
zwart []
kleur beslag <HS>
kleur <HF>
zwelling []
<HS> gezwollen
<HS> niet gezwollen
slijmvlies <HF>
omgevend mucosa <HF>
plooi <HF>
% van het aantal varices []
<HS> < 30%
<HS> 30 tot 60%
<HS> > 60%
<HS> alle varices
varices kleur <HF>
rode spots en streepjes op varices <HF>
1/2 van het lumen innemend (graad 3) []
grootte tov het lumen <HS>
1/4 van het lumen innemend (graad 2) []
grootte tov het lumen <HS>
30 tot 60% []
% van het aantal varices <HS>
< 30% []
% van het aantal varices <HS>
> 60% []
% van het aantal varices <HS>

Chapter 8

Summary, Conclusions, and Future Research

The aim of this study was to enhance the completeness and clarity of clinical narratives by supporting the capture of structured, coded data, with the challenge to approach the expressive power and flexibility of free text. As outlined in the introduction, we started our research with an *inventory of various aspects of current reporting*. This served as a basis for the *development and evaluation of a general model for structured data entry*. The inventory and evaluation were carried out in the endoscopy domain.

Inventory of Current Reporting

In Chapters 2 and 3, we assessed how well currently produced endoscopy reports reach their objectives: to serve as reference for follow-up examinations, and to communicate the findings of an endoscopic examination to the referring physician.

The inventory of endoscopy reports, described in *Chapter 2*, showed that these reports vary considerably in content and described detail. The results from the Delphi study, described in the same chapter, showed that endoscopists agree on the need for more information in their reports. Furthermore, we showed that endoscopists currently do not report what they believe should be reported. Although endoscopists differ with regard to which information they want to be added to reports, no evidence for fundamental differences in the opinions of the endoscopists was found. Therefore we concluded that an improvement of the contents of endoscopy reports is desirable and conceivable.

From the questionnaire among referring physicians, described in *Chapter 3*, we concluded that they also indicate a need for improvement. Whereas the endoscopists mainly criticized insufficiently described findings, referring physicians focus in their critique on inappropriate absence or presence of an indication, therapy, or follow-up plan, and on unclarity whether findings may account for complaints of the patient. Tailoring endoscopy reports to the needs of individual referring physicians will require explicit formulation of their expectations from an endoscopic examination.

In Chapters 4 and 5 we assessed ambiguity in terminology, and assessed interobserver variability in endoscopic descriptions.

The survey among endoscopists, described in *Chapter 4*, showed that there is a large overlap in the terms that endoscopists use to denote the size

of gastric ulcers. Since these terms are a potential for misunderstanding, we recommended the use of numerical size expressions.

The results of the study described in *Chapter 5*, showed that endoscopic descriptions of a given case differ per endoscopist, even when they describe the same topics and use the same set of terms. This descriptive inter-endoscopist variability is, again, caused by ambiguous use of terms. Reducing this variability will require that endoscopists assign the same meaning to terms, as well in words as in image. Including images with an endoscopy report will add to its value, but should not be regarded as a substitute for the written report: the textual interpretation reflects the personal view of an individual physician, and may explain the motivations for his decisions.

Thus, reporting in endoscopy needs to be improved, both in content and in terminology.

Development and Evaluation of the Descriptional Knowledge Model

We considered formalization to be the essential initial start towards improved reporting. No matter the format, more formal reporting will always produce tension between additional reporting effort and eventual benefit.

Therefore, we started *Chapter 6* by outlining the requirements for structured data entry (SDE) from both the users' and data usability point of view. These requirements led to the identification of descriptional knowledge as a specific type of knowledge that enables the definition of where and how concepts are used in clinical narratives. Linking concepts through relations and assigning them properties allows the creation of specific descriptional knowledgebases that contain the required descriptional knowledge for a given medical domain. Data capturing is made possible with a general data-entry program of which the behavior is determined by a combination of user input and the contents of the applied descriptional knowledgebase. The result of this structured data entry is semantically correct patient data in a structured coded format.

To obtain insight in the expressive power of our model, and the effects of the use of our formalism for reporting, we performed an evaluation, described in *Chapter 7*. From this study, in which we compared free-text

reports to reports made using our formalism, we concluded that the expressive power of our model is promising. Although we considered the uniformity of the acquired data for general, multipurpose usage not yet optimal, there was a beneficial gain with regard to the quantity and uniformity of these data: more subjects were more often described. Furthermore, ambiguity was reduced.

Thus, our proposed general descriptional knowledge model offers a viable potential for the support and improvement of medical reporting.

Discussion and Future Research

The research in this thesis describes a small initial step towards more formal medical reporting. The underlying assumption was that less ambiguous and more complete clinical narratives will benefit clinical care, and will be more suitable for research, decision support and quality assessment. Notwithstanding the promising results, the work in this thesis constitutes only a fraction of the work that lies ahead of its realization in daily clinical practice. In this last section, we therefore discuss topics that remained unanswered in this thesis, but that need to be addressed in the continuation of this research.

What are the expectations of our formalism when used for long-term routine reporting?

As set out in Chapter 7, our evaluation was carried out in an experimental setting. The results of that study are not necessarily representative for routine clinical use. Particularly because we believe that physicians will be motivated to record more data in a formal way, once they realize that they will also be able to benefit from the recorded structured data.

For routine use, however, it will be necessary to incorporate the formalism in the context of a computer-based patient record that has more general functionality (e.g., administrative functions). We also believe that routine use of our formalism will gain acceptability when the output will be in a more textual format instead of the current list of selected concepts. This would make reporting with our formalism more compatible to the present reporting process.

A key factor for acceptance will, furthermore, be its user interface; this

aspect has not received much attention during this research. Implementation of the formalism in a graphical environment will present new possibilities, and may give new insights for extensions of our model. For example, a graphical environment will enable to replace text menus by images, which will enable locations to be specified more readily. In the light of the low agreement between endoscopists on descriptive morphological features, as described in Chapter 5, it will be valuable to investigate the value of the addition of images that serve as a reference for descriptive terms.

In this thesis we made a distinction between techniques that a priori influence reporting and those techniques that take current free-text reports as input; natural language processing (NLP) techniques were categorized in the latter group. Although most groups working on NLP do not intervene in the reporting process, the underlying theoretical knowledge is complementary. We therefore believe that a combination of SDE and NLP techniques is feasible and deserves investigation, particularly to enhance the flexibility, and thus the acceptability, of data entry.

What are the expectations our formalism in other medical domains?

To address this question, we have recently been developing descriptional knowledgebases for general internal medicine (history and physical examination), pathology (cutaneous lymphomas), and radiology (chest X-rays). New domains may reveal new limitations of the formalism. Although the model has been extended since its status described in Chapter 6, we believe that additions to capture uncertainties, relations between findings, and temporal aspects deserve high priority. Since the evaluation in Chapter 7 showed that the use of 'normal' definitions led to a large increase of described data, further support of physician-specific definitions merits attention.

The factor that will predominantly determine the ultimate success of more formalized medical reporting is cooperation of physicians working in clinical practice. As long as we fail to demonstrate to clinicians the purpose and importance of more formal reporting, the tools provided by workers in medical informatics will remain unsuccessful. Since clinicians will be reluctant to accept enforced documentation of findings, we hope that the inventories described in this thesis will motivate them to improve their spontaneous reporting.

Finally, we feel that the importance of reporting is not sufficiently

stressed in medical education. Although medical students are taught which data are needed for their decisions, they are not taught when and how they should document such data. In our opinion, the described formalism may also be valuable as a reporting tool in medical education.

Samenvatting

Het doel van deze studie was het bevorderen van de volledigheid en duidelijkheid van medische verslaglegging door het ondersteunen van gestructureerde en gecodeerde invoer van gegevens. De uitdaging was hierbij de uitdrukkingskracht en flexibiliteit van vrije tekst te benaderen. Zoals vermeld in de introductie begonnen wij ons onderzoek met een *inventarisatie van verschillende aspecten van de huidige verslaglegging*. Dit diende als basis voor *de ontwikkeling en evaluatie van een algemeen model voor gestructureerde gegevens invoer*. De inventarisatie en evaluatie werden in het endoscopisch domein uitgevoerd.

Inventarisatie van Huidige Verslaglegging

In Hoofdstuk 2 en 3 werd bestudeerd in hoeverre huidige endoscopie verslagen voldoen aan hun tweeledige doel, namelijk het dienen als referentie bij vervolgonderzoek en het overbrengen van de bevindingen van een endoscopisch onderzoek aan de aanvragende arts.

De inventarisatie van endoscopie verslagen, beschreven in *Hoofdstuk 2*, liet zien dat deze verslagen aanzienlijk verschillen in inhoud en beschreven detail. De resultaten van de Delphi studie, beschreven in hetzelfde hoofdstuk, lieten zien dat endoscopisten meer informatie in hun verslagen nodig achtten. Verder bleek dat endoscopisten niet alles in hun eigen verslag documenteren wat zij noodzakelijk vinden. Hoewel endoscopisten verschillend antwoorden op de vraag welke informatie aan de verslagen toegevoegd moet worden, werd er geen aanwijzing gevonden dat die verschillen op fundamenteel andere meningen gebaseerd waren. Daarom was de conclusie dat een verbetering van de inhoud van endoscopie verslagen wenselijk en haalbaar is.

Uit de enquête, beschreven in *Hoofdstuk 3*, onder artsen die patiënten voor een endoscopie verwezen, concludeerden wij dat ook zij een noodzaak tot verbetering aangeven. Anders dan de endoscopisten, die voornamelijk onvoldoende beschreven bevindingen bekritiseerden, richtte de kritiek van de aanvragende artsen zich met name op ongewenste aan- of afwezigheid van een indicatie, een therapie of follow-up plan, en op onduidelijkheid over de relatie tussen de bevindingen en de klachten van de patiënt. Aanpassing van endoscopie verslagen aan de wensen van individuele verwijzende artsen maakt de expliciete formulering van hun verwachtingen van een endoscopisch onderzoek noodzakelijk.

In Hoofdstuk 4 en 5 bestudeerden wij dubbelzinnigheid in terminologie en interobserver variabiliteit in endoscopische beschrijvingen.

De enquête onder endoscopisten, beschreven in *Hoofdstuk 4*, liet een grote overlap zien in de betekenis van termen die endoscopisten gebruiken om de grootte van een ulcus pepticum te beschrijven. Omdat de gevonden overlap in de termen belangrijke klinische consequenties kan hebben, pleiten wij voor het gebruik van numerieke grootte aanduidingen.

De resultaten van de studie, die in *Hoofdstuk 5* wordt beschreven, lieten zien dat beschrijvingen van endoscopische bevindingen per endoscopist verschillen, zelfs wanneer zij dezelfde onderwerpen met dezelfde set aan termen beschrijven. Deze descriptieve inter-endoscopisten variabiliteit wordt, eveneens, veroorzaakt door dubbelzinnig gebruik van termen. Om deze variabiliteit te verminderen is het noodzakelijk dat endoscopisten dezelfde betekenis aan termen toekennen, zowel in woord als in beeld. Het toevoegen van beelden aan een endoscopie verslag zal de waarde van het verslag vergroten, maar het beeld moet niet als een vervanging van het geschreven verslag gezien worden: de tekstuele interpretatie reflecteert de persoonlijke kijk van een individuele arts, en kan de motivatie van zijn beslissing verklaren.

Dus, de verslaglegging in de endoscopie moet verbeterd worden, zowel in inhoud als in terminologie.

Ontwikkeling en Evaluatie van het Beschrijvings Kennis Model

Voor de verbetering van verslaglegging beschouwden wij formalisatie als essentieel. In welk formaat dan ook, formele verslaglegging zal altijd een spanning geven tussen de extra moeite van formeel verslagleggen en uiteindelijk voordeel.

Daarom werd *Hoofdstuk 6* begonnen met het uiteenzetten van de voorwaarden voor gestructureerde gegevens invoer, zowel vanuit de gebruiker als vanuit de bruikbaarheid van de gegevens. Deze voorwaarden leidden tot de identificatie van beschrijvings kennis als een speciaal soort kennis die het mogelijk maakt te definiëren waar en hoe begrippen in klinische beschrijvingen gebruikt kunnen worden. Het verbinden van begrippen door relaties en het toekennen van eigenschappen aan begrippen maakt het mogelijk specifieke beschrijvings kennisbanken te maken, die de benodigde beschrijvings kennis voor een bepaald medisch domein bevatten.

Het vastleggen van gegevens wordt mogelijk gemaakt door een algemeen invoerprogramma, waarvan het gedrag bepaald wordt door een combinatie van invoer door de gebruiker en de inhoud van de gebruikte beschrijvings kennisbank. Het resultaat van deze gegevens invoer zijn semantisch correcte patiënten gegevens in een gestructureerde en gecodeerde vorm.

Om inzicht te krijgen in de uitdrukkingskracht van ons model, en de effecten van het gebruik van ons formalisme op het verslagleggen, voerden wij een evaluatie uit, welke beschreven wordt in *Hoofdstuk 7*. Uit deze studie, waarin vrije tekst verslagen vergeleken werden met verslagen die met ons formalisme geproduceerd werden, concludeerden wij dat de uitdrukkingskracht van ons model veelbelovend is. Hoewel wij de uniformiteit van de verkregen gegevens voor algemeen gebruik suboptimaal vonden, was er wel een toename in de kwantiteit en uniformiteit van de gegevens: meer onderwerpen werden vaker beschreven. Ook waren de gegevens minder dubbelzinnig.

Dus, ons algemene beschrijvings kennismodel is een levensvatbare mogelijkheid voor het ondersteunen en verbeteren van medische verslaglegging.

Discussie en Toekomstig Onderzoek

Het onderzoek in dit proefschrift beschrijft slechts een kleine begin stap naar formele medische verslaglegging. De onderliggende aanname was dat minder dubbelzinnige en meer complete klinische verslagen de klinische zorg ten goede komen, en meer geschikt zullen zijn voor wetenschappelijk onderzoek, beslissingsondersteuning en kwaliteitsbewaking. Niettegenstaande de veelbelovende resultaten vormt het werk in dit proefschrift slechts een fractie van het werk dat voor ons ligt tot de realisatie van formele verslaglegging in de dagelijkse klinische praktijk. In deze laatste paragrafen bespreken wij daarom onderwerpen die onbeantwoord bleven in dit proefschrift, maar waaraan aandacht geschonken dient te worden in de voortgang van dit onderzoek.

Wat kunnen we verwachten van ons formalisme wanneer het gedurende langere tijd routinematig voor verslaglegging wordt gebruikt?

Onze evaluatie van het formalisme was, zoals vermeld in Hoofdstuk 7,

uitgevoerd in een experimentele setting. De resultaten van die studie zijn niet noodzakelijkerwijs representatief voor routinematig klinisch gebruik. Dit met name omdat wij geloven dat artsen pas dan gemotiveerd zullen zijn meer gegevens op een formele manier vast te leggen wanneer ze ook kunnen profiteren van de voordelen van vastgelegde gestructureerde gegevens.

Echter, voor routinematig gebruik zal het nodig zijn het formalisme te integreren met een elektronisch medisch dossier met meer algemene functionaliteit (b.v. administratieve functies). Verder zijn wij van mening dat het formalisme sneller geaccepteerd zal worden wanneer de uitvoer in een meer tekstueel formaat zal zijn dan de huidige uitvoer, namelijk een lijst van geselecteerde begrippen. Dit zou het verslagleggen met ons formalisme meer compatibel met het gangbare verslagleggings proces maken.

Voor acceptatie zal verder de gebruikers interface een sleutelrol spelen. Aan dit aspect werd tijdens dit onderzoek nog niet veel aandacht geschonken. Implementatie van ons formalisme in een grafische omgeving zal nieuwe mogelijkheden geven, en kan nieuwe inzichten geven voor uitbreidingen aan ons model. Een grafische omgeving geeft de mogelijkheid tekstuele keuzeschermen te vervangen door beelden, waardoor het selecteren van bijvoorbeeld lokaties eenvoudiger en directer wordt. Gezien de lage overeenstemming tussen endoscopisten over beschrijvende morfologische kenmerken, zoals beschreven in Hoofdstuk 5, verdient het aanbeveling de waarde van beelden, die als referentie kunnen dienen voor beschrijvende kenmerken, te bestuderen.

In dit proefschrift hebben we een onderscheid gemaakt tussen technieken die a priori het verslagleggen beïnvloeden, en technieken die huidige vrije tekst verslagen als invoer nemen: natuurlijke taalverwerking schaarden we bij de laatste categorie. Hoewel de meeste groepen die aan natuurlijke taalverwerking werken niet ingrijpen in het verslagleggingsproces, is de onderliggende theoretische kennis complementair. Wij geloven dan ook dat een combinatie van gestructureerde gegevens invoer en natuurlijke taalverwerking mogelijk is en aandacht verdient, met name om de flexibiliteit, en dus de acceptatie, van gegevens invoer te vergroten.

Wat kunnen we verwachten van ons formalisme in andere medische domeinen?

Om deze vraag te beantwoorden zijn we recentelijk begonnen

beschrijvings kennisbanken te ontwikkelen voor algemene interne geneeskunde (anamnese en lichamelijk onderzoek), pathologie (cutane lymfomen) en radiologie (thorax foto's). Nieuwe domeinen kunnen nieuwe beperkingen van ons formalisme aan het licht brengen. Hoewel het model al uitgebreid is sinds de status beschreven in Hoofdstuk 6, geloven wij dat toevoegingen voor het vastleggen van onzekerheden, relaties tussen bevindingen en temporele aspecten een hoge prioriteit hebben. Daar de evaluatie in Hoofdstuk 7 liet zien dat het gebruik van 'normaal' definities tot een grote toename in beschreven gegevens leidde, verdient het verder ondersteunen van arts-specifieke definities ook aandacht.

De factor welke voornamelijk het uiteindelijke succes van formele medische verslaglegging bepaalt is echter de medewerking van klinisch werkende artsen. Zolang wij niet in staat zijn de klinici het belang en de zin van formele medische verslaglegging te laten inzien, zullen de gereedschappen die door medisch informatici ontwikkeld worden zonder succes blijven. Daar klinici niet snel geneigd zullen zijn verplichte documentatie van bevindingen te accepteren, hopen wij dat de inventarisaties in dit proefschrift er aan bij zullen dragen dat artsen meer gemotiveerd worden hun spontane verslaglegging te verbeteren.

Als laatste menen wij dat het belang van verslaglegging nog immer onvoldoende benadrukt wordt in het medisch onderwijs. Hoewel aan medische studenten wordt geleerd welke gegevens van belang zijn voor het nemen van beslissingen, wordt hen onvoldoende geleerd wanneer en hoe die gegevens gedocumenteerd dienen te worden. Naar onze mening kan het beschreven formalisme ook waardevol zijn als onderwijs instrument.

Curriculum Vitae

Peter Willem Moorman was born in Diepenveen, the Netherlands, on the 10th of September 1963.

He received his undergraduate education at the Geert Groote College (VWO-B) in Deventer. He attended medical school at the Rijks Universiteit in Groningen.

During his medical studies he became interested in computers and informatics, and was introduced to the Medical Informatics discipline.

After receiving his MD in 1989 he worked for 6 months as a House Officer in surgery at the Royal Berkshire Hospital in Reading, England. This was followed by his military service.

In 1991 he started his PhD research in Rotterdam: a joint project of the Department of Medical Informatics of the Erasmus University, and the Department of Internal Medicine of the Dijkzigt University Hospital.

In 1995 he received his Master of Science in Medical Informatics.