Methods MIM 0081

© 2002 Schattauer GmbH

Text Generation in Clinical Medicine – a Review

D. Kraus
Heart and Diabetes Centre, Northrhine-Westphalia, Bad Oeynhausen, Germany

Summary

Objectives: This article aims at an analysis of ways of producing documents (such as findings or referral letters) in clinical medicine. Special emphasis is given to the question of whether the field of "Natural Language Generation" (NLG) can provide new approaches to ameliorate the current situation.

Methods: In order to assess the currently used techniques in text production, an analysis of commercially available systems was performed in addition to an extensive review of the literature. The sketch of current NLG approaches is also based on a literature review.

To estimate the applicability of several techniques to clinical documents, a typology of documents in clinical medicine was developed, based on rhetorical structure theory, speech act theory and certain recurrent linguistic phenomena exposed in the said documents.

Results: Current ways of producing text for documents in medicine are less than optimal in several respects. The field of NLG draws on the idea of generating text from a conceptual representation of not only certain facts, but also knowledge about how to express them via (written) language.

Unfortunately, NLG does not yet offer "ready-to-run" solutions for the automatic production of most of the document types in the given typology.

It seems, however, highly plausible that the demands of medical informatics for these kinds of systems will be satisfiable as NLG matures.

Conclusions: NLG offers a promising way of generating text for clinical documents, a problem of enormous economical importance. The medical informatics community should therefore commit itself to the idea of NLG in medicine.

Keywords

Natural language generation, clinical documents, document typology

Methods Inf Med 2003; 42: ■■■

Received February 18, 2002
Accepted June 7, 2002

1. Introduction

The task of dictating, for instance, referral letters is a stupefyingly repetitious one. In addition, the sheer workload which it imposes on medical personnel is of considerable economical importance. To illustrate: The German "Aktionsforum Telematik im Gesundheitswesen" (ATG) estimates that the equivalent in manpower of 125 Mio. Euro could be saved by reducing the average time required by a physician to produce a discharge letter by five minutes (1). The consideration that not only these letters, but also procedural reports (surgical, interventional etc.), findings and internal referral notes are typically produced by dictation and transcription, lends even greater importance to more efficient ways of generating these documents. The current approaches aimed at such a goal, i.e. structured forms, the use of "canned text" mechanisms, automatic generation employing mail merge techniques and speech recognition, are discussed in the first part of this review. Some of their failures can be traced back to linguistic features inherent in the modalities of language use prevalent in documents from the field of applied medicine. These linguistic phenomena, and the types of documents in which they prevail, are described in part two. Fortunately, there is a subfield of computational linguistics which for some 15 years has been addressing the problem of producing text automatically under the heading of "natural language generation" or NLG for short. With the exception of some research groups ((2-4) for instance), this field, its concepts and techniques seem to have received only part of the due attention from the medical informatics community. Medical informatics has therefore failed to pose the intricacies of the medical sublanguage as a problem worth the while of NLG researchers. In part three of this review I will attempt an account of the state of the art of NLG technology. The last section sets out to reconcile the pressing need for automatic generation of clinical documents with the sometimes and – in this context – too ambitious goals of current NLG research.

2. How Do We Do It Today?

With the advent of information and communication technology (ICT) in the working field of the physician, some alternative ways of producing clinical documents have emerged. Ignoring the case of letting the physician type the text herself, there are four main techniques for reaching this goal. They are realized in off-the-shelf systems for documenting examinations (e.g. EndoBase™ (5)), practice computer systems (such as Medistar™ (6)) or hospital information systems (ORBIS Med™ (7), for instance) or as speech recognition solutions with specialized dictionaries for medicine ((8, 9)).

2.1 Structured Data Entry (SDE)

The idea of SDE (see (10-16)) in this context is to provide the physician with predefined forms, containing the typical graphical user interface (GUI) widgets (such as checkboxes, radio buttons, popup menus etc.), and to build a structured report from the data entered. This approach is sufficient as long as

- An agreed nomenclature can be utilized
- The resulting documents' uniform structure is acceptable
- The necessarily telegraphic (and often clumsy) style of output text is tolerated.

These conditions being fulfilled, structured data entry has some notable advantages:

- A normative force, i.e. the use of certain formulations can be enforced or suppressed.
- An educative value, since less experienced physicians are offered a guideline to the complete account of all observations and/or facts.
- Plausibility and completeness checks can be utilized.
- Structured data already present can be integrated, thus reducing the total effort on the physician's side.
- The data can be used for other purposes too, such as quality assurance or scientific studies.

As a consequence, there were, and still are, a couple of systems (17-20) in routine use which are based on this approach. However, the premises mentioned above are not always given and whereas an ECG finding in "telegraphese" may be perfectly acceptable, quite the opposite may hold for a discharge letter after complicated heart surgery. Two other drawbacks of the SDE/ report approach can be identified:

Stylistic variance is not possible. The rigid one-to-one mapping of a user-entered data item to a phrase in the report not only yields less than natural text, but also leaves no place for the stylistic preferences of different physicians. What may be a desirable quality in one context, ensuring standard terminology, can be a serious obstacle to broad user acceptance under circumstances where physicians do not want their individual formulations to become streamlined by an ICT system.

The (temporal) decoupling of data entry and report generation which many of the systems exhibit (e.g. (20, 21) or (14)) is a step back from a "what you see is what you get" functionality. Especially in the early phases of system use, it can be irritating for a physician not to have an immediate response regarding what his action in the GUI will result in in the output text. Naturally enough, the user will have less confidence in the report being a faithful account of his entries the longer it takes until he has the possibility of reviewing it.

2.2 "Mail Merge" Generation from Databases

Typically found in contexts where a number of data items are stored in a database, such as laboratory systems, patient data management systems (PDMS) in intensive care units, but also in conjunction with SDE forms, the technology of mail merge features can be used to build up a text. With procedural constructs such as "IF…THEN…ELSE", "CASE", "LOOP" etc. it is possible to put together quite complicated text structures (this approach has been adopted for instance in the PHOENIX-system (22)). Nevertheless, an enormous amount of coding has to be carried out just to get the syntactical structure right, at least in languages where there is a high degree of flexion in nouns, pronouns, adjectives and verbs. Even though more would be possible in theory, in reality these systems tend to adopt an only slightly less telegraphic style and are thus subject to the same criticisms as the SDE/ "concatenated item labels" approach.

2.3 "Canned Text"

Whereas the techniques described in 2.1 and 2.2 required a system analysis/design phase before actual use and in most cases are configurable only by the ICT vendor, the implementation of canned text phrases is very easy, and they are a commodity item in virtually every text processing environment. The user has total control over the inventory of text phrases. This is, of course, a blessing in disguise, as the full responsibility for completeness, stylistic adequacy and syntactical congruence of the phrases is imposed on the system user. Thus, in many cases only normal findings are represented as canned text phrases. Wherever complicated sentence structures must be built from several variable text fragments, it is more likely than not that, at least in languages with much inflexion, the number of phrases suffers from combinatorial explosion. This again can be avoided only by adopting an agrammatical "telegraphic" style (a commercially available system serving as an example of this approach is OrthoStar (21)). It is worth noting that, when this is done, the canned text approach has most of the drawbacks of the other two approaches, but none of their advantages, rendering it the least favorable alternative!

Of course a lot of valuable work has been done on the basis of the techniques mentioned above. Recognizing the need for flexible and open, self-documenting and communicable reports, Kahn (23) for instance has dedicated much effort to the task of platform-independent reporting. But irrespective of the accomplishments in medical linguistics, exemplified for instance in the Unified Medical Language System (UMLS) (24, 25), the technology for generating the appropriate text from these reports is still lacking.

2.4 Speech Recognition

Speech recognition technology is evolving fast, and it is a safe assumption that it will contribute significantly to the production of textual documents in health care (see (26)). This contribution is likely to become even greater when speech recognition is not used in isolation, but is combined with other approaches. Canned text phrases can, for instance, be inserted into a document by spoken commands. Very interesting in this context is the approach of Nugues et al ((27)): They combine speech recognition with a definite clause grammar formalism ((28)) into a dialog system where the physician builds up a medical report by answering questions orally.

Nevertheless, in addition to its well-known problems, such as the need to build up a dictionary, the necessity of a training phase, sensitivity to alterations in speech modulation, problems with non-native speakers, etc., speech recognition used in isolation lacks some of the advantages of the other approaches. Neither the norma-

tive or educational value, nor the possibility of integrating preexisting data or reusing the resulting text in other contexts can be realized with speech recognition. The scenario of a physician re-dictating a paragraph from, say, an endoscopic finding as part of a discharge summary is still a far cry from notably decreasing that physician's workload.

Taking the pros and cons of the approaches outlined above, one could be tempted to formulate an "ideal" system:
It should generate

- Concise, but nevertheless "natural" text
- with varying complexity and according to different stylistic preferences
- from preexisting data and/or SDE input
- with instant feedback.

Moreover, such a system should be adaptable to different medical specialties, different document types and different ICT environments.

These are admittedly high demands, and they raise the question of whether such a system is possible at all. A closer examination of the linguistic characteristics of textual medical documents seems required.

3. A Typology of Documents in Clinical Medicine[1]

In order to answer the above question, one needs to know which specific linguistic properties of documents in clinical medicine have to be reproduced for a generated text to look "natural". Such an analysis can be twofold: Firstly, there are certain linguistic phenomena which are mainly employed to ensure the conciseness and cohesion/coherence of the documents. These are (explanation follows) "aggregation", "anaphora" and the use of "referential expressions". In addition, there is an abundant use of "ellipsis" and "enumerative phrases".

The other way of looking at the documents at hand is to analyze their internal structure, the "text organization". An instrument for doing this can be derived from the rhetorical structure theory (RST, (29-33)), a widely accepted (and augmented/modified) framework for identifying the relations which can hold between the parts of an utterance.

3.1 Linguistic Phenomena

3.1.1 Aggregation

Aggregation is a very broad term which has been divided into several subtypes, such as "conceptual", "syntactical", "lexical aggregation" etc. (see (34-39)). Some examples of these different subtypes may suffice to illustrate the concept, which can be roughly described as a linguistic means of "redundancy elimination".

"Conceptual aggregation" denotes the process of combining several terms into just one, which encompasses and implies them. The phrases "edema of the left ankle" and "edema of the right ankle" can be aggregated (quite trivially) to "edema of both ankles". Similarly, but far less obviously, "aplastic anemia", "thrombopenia" and "agranulocytosis" can be aggregated into "panmyelopathy". In fact, performing conceptual aggregation may require a great deal of domain knowledge and can be construed as drawing inferences from a knowledge base. Its effect is a more concise, readable text using the adequate terminology. Omitting it (for instance in canned text approaches) may lead to text which is not only unnatural, but often inappropriate and sometimes even inconceivable.

"Syntactical aggregation" denotes the coordination of sentences which are (implicitly or explicitly) connected by "and". So the sentences "The patient has hypertension" and "The patient has angina" can be aggregated to "The patient has hypertension and angina". While this seems an easy thing to do at first sight, there are many cases where exactly the same operation leads to undesirable or even wrong results: "The vessel is prepared from myocardial tissue" and "The vessel is mildly sclerotic" would yield "The vessel is prepared from myocardial tissue and mildly sclerotic", which sounds awkward. And aggregating the last two sentences in "Preparation of the LCA. It is entirely sclerotic. It is improbable that the patient would benefit from a bypass" to give "Preparation of the LCA. It is entirely sclerotic and improbable that the patient would benefit from a bypass" would be downright wrong. There are, moreover, cases of coordination which are based on a different case of aggregation, e.g.: "m.biceps and m.triceps are atrophic" can be regarded as aggregated from "m.biceps is atrophic" and "m.triceps is atrophic", but "m.biceps and m.triceps are antagonistic" is quite a different thing. So-called semantic aggregation would be the process of producing this sentence from "m.biceps is antagonistic to m.triceps" and "m.triceps is antagonistic to m.biceps".

3.1.2 Anaphora

Terms which stand for entities or circumstances mentioned in the preceding text are called anaphora. So in "We prepare the LCA from myocardial tissue. The vessel is mildly sclerotic. It is incized and probed.", the term "the vessel" is an anaphor to "the LCA", and "It" is anaphoric to "the vessel". Again, failure to employ anaphora may lead to clumsy and unnatural text such as "We prepare the LCA from myocardial tissue. The LCA is mildly sclerotic. The LCA is incized and probed.". It is important to recognize that not only nouns and pronouns can be used in such a manner. In "We open the pericardium. There we find…" the "there" may be construed as referring anaphorically to the entire cardiac site which has been introduced only implicitly.

3.1.3 Referential Expressions

Anaphora represent only one of the many ways the human language (and medical sublanguage) may refer to entities. One may freely exchange synonyms (such as LCA and RIVA), one can flesh out nouns with a number of attributes and/or relative clauses as in: "The slightly elongated and varicose vein which had been prepared

[1] Acronyms used in this section: "LCA"-"left coronary artery" "RIVA" – "ramus interventricularis anterior", the latin term for LCA"m." as in "m. biceps"- "musculus" = "muscle"

from the left lower leg is used to…". Expressions may refer to entities introduced later in the text (kataphoric expressions like the "it" in "It's a large solid tumor that can be palpated in the left lower abdomen"). References can be made to certain points in time or space and to single persons or groups (temporal deixis: "Then we did…", spatial deixis: "Here we found…", personal deixis: "The patient's mother had…". Being a broader concept than anaphora, "referential expressions" in general contribute to the "connectedness" of a text and thus again to its readability and naturalness.

3.1.4 Ellipsis

Sentences where phrases have been omitted are called *elliptic*. In "Male white Caucasian, normal weight, no history of smoking or drug abuse." there is no verb and the sentence is therefore elliptic, but it may well be acceptable. For several reasons, dictation and reading efficiency being obvious ones, ellipsis is a very prominent feature in medical documents.

3.1.5 Enumerations

Quite often, aggregation leads to sentences in which a number of entities or features are referred to in an enumerative manner: "The prepared vessel is not sclerotic or varicose, it is sufficiently long, has a wide lumen and a good flow." is the shortest way to make five propositions about a single item. But the linguistic construction of these enumerations is far from trivial. Joining all the items but the last with a comma and inserting an "and" before the last one works only for the minority of cases. Complications arise from

- the grammatical nature of the items, as noun phrases, adjectives, adverbs and even verbal phrases may be enumerated. Of course, categories may also be mixed, as in the above example.
- the polarity of the enumeration. In the negative case we find "neither-nor" constructs instead of the ones with "and", and polarities can be mixed as well, giving an "a, b and c, but neither d nor e" structure.
- the modality of the enumerative construct: It can relate a temporal succession ("firstly a, then b and c, finally d") a spatial coverage of a region or an exhaustive enumeration of a conceptual set. A description of coronary status is, for instance, likely to be an enumeration of propositions about the set of coronary arteries.
- the context in which the enumeration occurs. Especially the ordering of the items may depend on preceding or following text.

As is the case with ellipsis, enumerations are powerful instruments, enabling short and concise text. Failure to generate valid enumerations can render the text not only overly verbose but even inacceptable. As an example of the latter consider the slightly modified previous example in its context: "The prepared vessel is not sclerotic or varicose, it is sufficiently long, has a wide lumen but very poor flow. We therefore discarded it as a CABG graft." It is easy to see that any ordering of the five attributes where the "poor flow" expression does not come last leads to a problem with the ensuing "We therefore…".

3.2 Document Structure

Another viewpoint from which the linguistic features of medical documents can be analyzed is, as mentioned already, the "content structure" of a text. The rationale of such a text-organizational approach stems from the fact that the conceptual relations between segments of a text may have an influence on its surface structure and that, in order to build a (medical) document, this conceptual structure must be represented in the generating system as well.

Several formalisms are suitable for such a study (for example "centering" (40), "discourse representation theory (41), the abovementioned rhetorical structure theory or even "speech act theory" (42-45), for an overview see (46) or (47)).

The key idea behind rhetorical structure theory, which is chosen here as an example, is the identification of the relations which can hold between two arbitrary linguistic constructs (termed "nucleus" and "satellite"). Although no standard body of "rhetorical relations" exists, some of them are fairly common. They include:

"volitional result" – the nucleus (N) is the result of an intentional action or planned process (satellite, S), as in "the circulatory situation could be stabilized (N) by the admistration of drug X (S)"

"non-volitional result"– the nucleus is the result of an unintended action or unplanned process (satellite), as in "EEG activity was obliterated for 45 minutes (N) by the increased use of hypnotics.(S)"

"means" – the nucleus denotes the means or instrument to achieve a certain state of affairs (S), as in "patient was referred for an echocardiogram (N) to further determine the anatomic features of the ventricular septal defect (S)".

"*concessive*" – the nucleus is taken for granted, even in the light of other information (satellite), as in "The possibility of Devic's disease as a presenting manifestation of multiple sclerosis cannot be ruled out, even though CSF oligoclonal bands were negative."

"background" – the nucleus denotes background or circumstantial information. Example: "The 75-year old female patient had a history of chronic severe hypertension and transient confusion (N) when she underwent revascularization of three coronary arteries."

Other relations typically to be found in documents from the domain of clinical medicine are: "contrastive", "elaboration", "purpose", "motivation", "logical condition" and "enablement".

Using these relations, a hierarchical decomposition of documents can be given. The relevance of such an analysis for text generation is quite obvious: It lies in the fact that these relations are regularly signaled by linguistic means. In particular, "cue words", such as "although", "thus", "because", "to this end" or "for instance", are strong indicators of rhetorical relations. A system for generating documents in which certain rhetorical relations hold must clearly be able to generate the appropriate cue words.

3.3 The Typology

The following is a typology based solely on the linguistic complexity of the respective documents, completely disregarding non-linguistic features.

3.3.1 C4: Lab Results

The simplest case of a textual document is the laboratory finding: To the extent that it contains more than numerical values, unit descriptors and markers for pathological parameter values, it is a mere enumeration of simple (elliptic) phrases.

3.3.2 C3: Routine Examination Findings/ Clinical Notes

Fairly standardized findings of technical routine examinations (for instance: ECG findings) fall into the next category: Their descriptive part is very often a plain enumeration of single phrases, a little aggregation may occur (as in "in all ECG leads"). Additionally, in the interpretative part, a limited number of rhetorical relations are to be found (concessive, elaboration, background). Since none of the means to enforce text cohesion/coherence are used, these documents mostly use a telegraphic style. Clinical notes in a patient's chart, in a very similar way relating loosely coupled observations, fall into the same category with regard to their linguistic complexity. The most prevalent rhetorical relations here are "(non)volitional result", "motivation/purpose".

3.3.3 C2: Complex Examination Findings/ Interventional and Surgical Reports

The less standardized, the less routine and the more complex examinations are, the more linguistic phenomena can be found in the findings. The reason for this can be derived from the fact that procedural details, deviation from routine protocols and interpretation of ambiguous findings all have to be accounted for explicitly, i.e. the text becomes more argumentative. Reports of interventional procedures or surgery are among the most complicated items within this category: They display nearly all the rhetorical relations mentioned: aggregation, referential expressions including temporal and spatial deixis, anaphora and, of course, enumeration.

3.3.4 C1: Discharge Letters

Discharge letters form the next category, but their separation from C2 requires methodological equipment additional to rhetorical relations. Due to space restrictions this is not elaborated further, but merely hinted at: When analyzing clinical documents with the instrument of speech act theory[2] one finds that discharge letters are the only documents in which illocutionary acts play an important role. Especially directive and commissive acts are quite common in discharge letters.

3.3.5 C0: Medical Expert Opinions

These constitute the most elaborate category of documents in clinical medicine with regard to text generation. They display all the complex linguistic features of discharge letters, yet they differ from them in two ways: Their subject is not only the patient case, but also the current body of scientific knowledge and accepted state of the art. Secondly, they may adopt a "meta"-perspective with respect to other clinical documents, not simply relating them but talking about them, their presence or absence, their validity and the like. This of course has implications for textual realization (indirect speech, for instance) to the effect that expert opinions differ not only in domain but also in linguistic diversity and complexity from more routinely produced documents.

4. Natural Language Generation

For the purpose of this review we will regard as an NLG system any system which technologically goes beyond the approach of filling strings into a template to generate text from an underlying representation. The first attempts in this direction date back to the early 1970s: The SHRDLU system, for instance, produced simple English sentences from an underlying representation. Other systems with differing goals, architectures, methods and, of course, success have followed (see (48)). For the last 15 years, NLG seems to have been in a sort of consolidation: The first system in routine use was FOG (49), there are SIGs, regular workshops and conferences and even textbooks which successfully attempt to present the current state of the art (50) (for an overview see also (33) or (51)). One major undertaking has been an effort to develop a standard reference architecture for generation systems (RAGS (52)). For the subtask of "surface generation", which is the process of generating final text based on an abstract syntactical specification, two quite elaborate systems have been developed and used in academic projects, namely *FUF/SURGE* (53, 54) and *KPML* (formerly *PENMAN*, (55)). *RealPro* (56, 57), another system for surface realization, is commercially available. The list of NLG systems (48), which is constantly being updated, contains more than 100 entries. On the other hand, the field has not yet matured to a point where there is an established body of concepts and techniques which could be applied directly to most NLG problems. Nevertheless, a short sketch of NLG concepts and their application in the health care environment seems worthwhile.

4.1 NLG Concepts

A hypothetical "universal" NLG system would perform the following tasks:

- a) discourse planning
- b) content determination, document structuring
- c) multilingual microplanning with lexicalization, aggregation, generation of deictical and other referential expressions (ellipses etc.)
- d) (multilingual) morphosyntactical realization,
- e) formatting

[2] The terms "illocutionary act", "directive act" and "commissive act" come from speech act theory

Discourse planning is necessary wherever an NLG system is not producing a single textual utterance but takes part in a dialogue. Of course, the entire problem field of building and maintaining user models emerges here, as does the problem of planning (for an overview see (46)). The next task, content determination, is dedicated to the problem of "what to say" in an utterance (or document, in our case). While this may be trivial in some cases, it is a hard problem to solve in others. One only needs to imagine the process of generating a discharge letter based on a comprehensive clinical database to realize the severity of the problem.

The next "task", C), looks like an unstructured enumeration of diverse subtasks, and that is exactly what it is. There is little agreement about which subtask belongs where. Some researchers (58) even deny that aggregation could be a subtask (or several subtasks, for that matter) in the first place. They argue that aggregation is ubiquitous, can occur in countless subtasks and should thus be rather construed as an emergent phenomenon. At any rate, lexicalization (the process of selecting lexical items for conceptually represented entities), as well as the other subtasks of microplanning, have to be performed before morphosyntactical realization, the ultimate generation of sentential structures containing inflected forms, takes place. A final phase of formatting may be mandatory in some cases.

As already hinted at in the subtask names, multilinguality can be an important goal for NLG. The TECHDOC system (59), for instance, was aimed at the multilingual generation of technical documents. The work of Wagner et al. (2, 60, 61) was directed at a multilingual generation of descriptive noun phrases for medical procedures in the context of GALEN-IN-USE (62). Other goals pursued in the construction of NLG systems are:

- correct generation of referring expressions (63), including quantified (64), "distributed" (65) and vague (66) descriptions. These include furthermore the generation of different types of anaphora ((67-69)), especially "one"-anaphora (70) and personal/temporal deicticals (71)
- multimodality: integration of NLG techniques within hypertext (72, 73), graphical output (74) and gestures (75).
- stylistic variance: the style of the generated text, including but not restricted to its level of formality and verbosity, use of technical terms, grammaticality etc., depends on its intended use, expected readers, external guidelines, rules or even laws. Patient information leaflets differ notably in style, for instance, from a referral letter, even when the content is identical.

4.2 NLG Applications in Medicine

Cawsey et al. in (4) give a short overview of NLG applications in health care. They isolate as principal functionalities:

- producing explanations and giving advice in medical expert systems
- generating reports, progress notes and discharge letters
- preparation of individualized patient information material
- generation of descriptions of medical concepts.

A good example of the last category is the work of Wagner et al. mentioned above.

In the case of personalized patient information, the work of Reiter et al. (76) on STOP, a system writing individualized leaflets to motivate patients to stop smoking, is a well-known example. PIGLIT (77) generated hypertextual information for diabetes patients. Being linked to the patients' electronic records, it facilitated an exploration of the terms and concepts mentioned there. N. Elhadad and Mc Keown developed a system able to summarize medical articles specific to a patient (78). An earlier system by McKeown and coworkers, MAGIC ((79)), produced patient briefings on an intensive care unit.

In the area of generating reports and letters to health care professionals, a few systems can be named:

The earliest attempts can be traced back to the mid-1980s, for instance the Linguistic String Parser Project (80) or the BAIK system (81) which used the "canned text-mail merge"-technique. In the following years some researchers (e.g. (82, 83)) addressed the problem of generating medical documents in an ad-hoc, though not necessarily unsuccessful way, whereas others realized the need to draw on an underlying conceptual representation: Bernauer (84, 85) utilized conceptual graphs as a representational formalism from which reports (of bone scans) were generated. At approximately the same time, A. Rector, B. Nowlan et al. even addressed the multilinguality issue with PEN&PAD (e.g. (86)), and in 1994 Bullock (87) used the semantic net underlying PEN&PAD as input for a text generation component. It is worth recognizing that whilst the problem of data entry, i.e. achieving a situation in which physicians willingly enter data on their own, was not solved satisfactorily in all of these projects, their failure to come into widespread routine use must not be attributed solely to the drawbacks of the named approaches. The low quality of the resulting texts, especially with respect to their diversity, appropriateness and readability, in conjunction with their limited to non-existent potential to "tune" their generation component to individual needs or the progress of clinical medicine, must be regarded as being of at least equal weight. It is thus surprising that in (4) Cawsey et al. remark that for medical reports "simple methods can be used for realization, as frequently only a small number of sentence forms are needed". They nevertheless conceded that in the case of bone scan studies "the variety of possible kinds of observations requires a relatively sophisticated realization module", so one may freely draw the conclusion that this holds for every kind of report in which the variability of observations equals or exceeds the one of bone scan studies (which should be a fairly high proportion of the reports in a medical record). The fact that text quality really is an issue may be seen from the fact that although structured data entry (sometimes even based on conceptual representations) is becoming more and more common in medicine (11-14, 19, 88-90) there is still no system in routine use which

- for a nontrivial domain
- based on NLG principles
- generates medical reports with fluent, concise and readable text.

The reasons for this deplorable fact have to be found.

5. Does NLG Try to Solve the Wrong Problem?

One may indeed say that NLG does not address all the problems surrounding the intended application of medical document generation, but this is of course also due to the neglect of the medical informatics community. Moreover, NLG still has a long way to go before it can be expected to offer "shake-and-bake" solutions to every generation problem, so that the very idea of employing NLG to generate medical notes, reports and letters may be premature. On the other hand, the body of concepts and techniques already elaborated under the heading of NLG leads the way to far more elaborate systems for generating text in clinical medicine:

- The key idea is to abandon the handling of text chunks in favor of generation from syntactical representations.
- "Off-the-shelf" systems exist for the task of surface generation (at least for some important languages (53, 55)); they should be used wherever possible.
- The conceptual apparatus of NLG enables identification of the linguistic features which are essential for natural and concise text in documents of clinical medicine: Aggregation, referential expressions including anaphora and deicticals, enumerative phrases, ellipsis and linguistic markers for the rhetorical relations (38, 46, 91). This insight may provide a guideline for developing the functionality of systems.
- Analysis of the documents under consideration from an "NLG perspective" reveals (not surprisingly) that even a representation of semantics and pragmatics is mandatory in some cases, thus necessitating an ontology of the domain (either ad-hoc and for generation purposes only, or based on a reusable one already elaborated). Again, researchers have addressed knowledge representation and semantical issues in the NLG context (92). Especially Ceusters et al. ((93)) argue convincingly that given ontologies must be augmented when there are to be used linguistically.
- As for example the task of generating a referral letter completely "from scratch" without compromising the stylistic quality seems very hard to accomplish, it may well be the case that a mixture of more rigid textual templates with pure syntactical generation is called for. Research into this kind of hybrid system has already been undertaken (94, 95).

Thus, formulating a kind of "wish list" to guide further developments in NLG would be the right thing for the medical informatics community to do. The rationale behind, and justification for, such a list stems from the assumption that it could be fruitful to concentrate on the problems of a specialized application, such as medical reports. This domain has its special intricacies but, on the other hand, lacks some of the problems other domains expose. For instance, the entire area of discourse modeling and document structuring collapses into the application of simple "document skeletons". Many syntactical constructs never occur in medical reports or letters: Questions or commands are, wherever they appear, formulated in the indicative mood. Whereas current research in NLG is directed at single phenomena on the one hand or general architectural issues on the other, it may be the case that addressing and solving the text generation problems of a subdomain would not only benefit this domain, but could also prove to have the right level of generality to become the basis of usable systems within a reasonable time span.

From the author's point of view, the following items should be on the abovementioned wish list:

- Research into the peculiarities of the medical sublanguage, inasmuch as they may be relevant to text generation. For instance, enumerative constructs which are so abundant in medical documents have received little attention as such. The same holds for "ad hoc neologisms", that is the formation of new words which are nevertheless understood instantly by every health care professional.
- Intradocument variability, the use of different lexems and syntactic constructs for similar contents, is a particularly important feature for the acceptability of automatically generated text, but appears to have been outside the focus of NLG research until now.
- Work on "off-the-shelf" surface generators, specialized to the medical sublanguage and with built-in support for phenomena such as enumerations, conceptual aggregation and referential expressions.
- A dedication to hybrid systems, able to use both predefined templates and "from-scratch" generation. This seems a promising approach in order to get systems up and running within a shorter time frame, leaving the option open for later refinement.
- An evaluation methodology for automatically generated reports. Evaluation of NLG systems is a much discussed topic (96), and here again a restriction to application subdomains should prove useful.

The medical informatics agenda remains full. For instance, work on reusable ontologies (such as (62)) should explicitly take into account the issues of text generation. The same holds for efforts towards creating terminological systems (97-100). Structured data entry, being the main source of text generation, not only has to come into more widespread use, but must evidently e based on the same ontology as the generation component, a reuse which the ORCA approach (12), for instance, does not account for. This can be seen as a revival of the "knowledge acquisition bottleneck", which has haunted the Artificial Intelligence community for a long time.

6. Conclusions

In two recent papers concerning the future role of medical informatics, both Haux (101) and van der Lei (102) identify as central to this discipline several applications which make use of a comprehensive medical record, such as integrated decision support systems, observational databases or

health reporting. In the light of its indisputable economic significance, one is tempted to add to the list of these important targets the problem of generating text for medical reports, clinical notes, discharge summaries and referral letters from the medical record. This claim may be substantiated by the fact that this problem has been addressed by a great number of researchers over a period of more than 15 years, from early works by O'Keefe (103), Giere (81) and Li (80), the PEN&PAD project (86) and Bernauer's work on conceptual graph formalism for the generation of findings (84, 85) or Kuhn's work on gastroenterological reporting (104), to current approaches which have been at least partially influenced by NLG approaches (2, 3, 18, 105).

Nevertheless, there are still no "ready-to-use" systems for this task. The field of natural language generation, which has this as one of its goals, has the potential to come up with solutions, but much work remains to be done on both sides - medical informatics and NLG - before the enormous benefit which the automatic production of these documents would provide, can be realized. The "wish list" can be seen as a starting point. For both sides it would be worth the effort.

Acknowledgments
I would like to acknowledge the motivation, time and experience which Reinhold Haux dedicated to my work and especially this paper. Ehud Reiter's comments on earlier versions of this article have been stimulating and encouraging.

References

1. Hons J. Der elektronische Arztbrief ist ein Stück näher gerückt. Ärztezeitung 2001 5/2001 (3); Sect. 21.
2. Wagner JC, Rogers JE, Baud RH, Scherrer J-R. Natural language generation of surgical procedures. Int J Med Inform 1999; 53: 175-92.
3. Hahn U, Romacker M, Schulz S. Discourse structure in medical reports - Watch out! The generation of referentially coherent and valid text knowledge bases in the medSyndikate system. Int J Med Inform 1999; 53: 1-28.
4. Cawsey AJ, Webber BL, Jones RB. Natural language generation in health care. J Am Med Inform Assoc 1997; 4 (6): 473-82.
5. Olympus. ENDOBASE; 2001. www.olympus-europa.nl/endo_int/products/endobase/features.stm
6. Medistar. Medistar; 2001. www.medistar.de
7. GWI. ORBIS; 2001. www.orbis-net.de/
8. DragonSystems. Dragon NaturallySpeaking Medical V6.0. In; 2002.
9. ViaVoice. ViaVoice for Windows. In: Release 9 ed: IBM; 2002.
10. Lussier YA, Maksud M, Desruisseaux B, Yale P-P, St-Arneault R. Pure MD. In: A computerized patient record software for direct data entry by physicians using a keyboard-free pen-based portable computer. Frisse ME, ed. SCAMC 92; 1992; Baltimore: McGraw-Hill; 1992. p. 261-3.
11. Heathfield HA, Kirby J, Hardiker N. Data entry in computer-based care planning. In: Medical Informatics in Europe 94 (MIE 94); 1994; 1994. p. 186-9.
12. Ginneken AMv. Structured data entry in ORCA: the strengths of the two models combined. In: 20th Symposium on Computer Applications Med Care (SCAMC 96); 1996: Hanley & Belfus,Philadelphia; 1996. p. 767-801.
13. Colburn C. Structured text-documentation meets technology. J AHIMA 1997; 2.
14. Kuhn K, Zemmler T, Reichert M, Heinlein C, Roesner D. Structured data collection and knowledge-based user guidance for abdominal ultrasound reporting. In: Proceedings Ann Symposium on Comput Applied Med Care; 1993; p. 311-5.
15. Moorman PW, Gineken AMv, Lei Jvd, Bemmel JHv. A model for structured data entry based on explicit descriptional knowledge. Methods Inf Med 1994;33:454-63.
16. Yamazaki S, Satomura Y. Standard method for describing an electronic patient record template: Application of XML to share domain knowledge. Methods Inf Med 2000; 39: 50-5.
17. Hohnloser JH, Engelmeier T. Dateneingabe durch Ärzte in elektronische Patientenakten: Vergleich zweier Methoden im 5-jährigen klinischen Routinebetrieb. In: GMDS 96; 1996.
18. Kraus D. Suregen2 – a model based generator for surgical reports. In: MIE 2000. IOS Press; 2000.
19. DeFriece RJ. MedIO: A program for intelligent clinical data entry. In: 19th Annual Symposium on Computer Applications in Medical Care (SCAMC 95). Gardner RM, ed. New Orleans, Louisiana, USA: Hanley & Belfus; 1995. p. 1005.
20. Zacher M. Computergestützte Befunddokumentation in der Arthrosonographie (M. D.). Giessen: Justus-Liebig-Universität Giessen; 1994.
21. Veigel B. OrthoStar; 2001. www.medistar.de/sites/fachaerzte/ orthostar.html
22. AG P. Phoenix; 2001. www.phoenix.ch
23. Kahn CE. A generalized language for platform-independent structured reporting. Methods Inf Med 1997;36:163-71.
24. McCray AT, Hole WT. The scope and structure of the first version of the UMLS semantic net-work. In: Fourteenth Annual Symposium on Computer Applications in MedicalCare (SCAMC-90). Miller RA, ed. Washington DC: IEEE Computer Society Press; 1990. p. 126-33.
25. McCray AT, Nelson SJ. The representation of meaning in the UMLS. Methods Inf Med 1995;34(1-2):193-201.
26. Mönnich G, Wetter T. Requirements for speech recognition to support medical documentation. Methods Inf Med 2000; 39 (1): 63-9.
27. Nugues N, Godéreaux C, Guedj POE, Cazenave F. Question answering in an oral dialogue system. In: Annual International Conference of the IEEE Engineering Medicine Biology Society; 1993; p. 590-1.
28. Pereira FC, Warren DH. Definite clause grammars for natural language analysis. Artificial Intelligence 1980;13.
29. Mann WC, Thompson SA. Rhetorical structure theory: A theory of text organisation. Technical Report. Marina del Rey, CA, USA: USC/Information Sciences Institute; 1987. Report No.: RS-87-190.
30. Moore JD, Pollack ME. A problem for RST: The need for multi-level discourse analysis. Computational Linguistics 1992;18(4):537-44.
31. Rosner D, Stede M. Customizing RST for the automatic production of technical manuals. In: Sixth International Workshop on Natural Language Generation; 1992; Duisburg: Springer Verlag, Heidelberg; 1992. p. 199-215.
32. Dalianis H, Johannesson P. Explaining conceptual models - using toulmin's argumentation model and RST. In: Third International Workshop on the language Action Perspective on Communication Modelling (LAP 98); 1998; Stockholm, Schweden; 1998. p. 131-40.
33. Cole RA, Mariani J, Uszkoreit H, Zaenen A, Zue V. Survey of the State of the Art in human language technology; 1996. www.cslu.cse.ogi.edu/HLTsurvey/HLTSurvey.html
34. Shaw J. Conciseness through aggregation in text generation. In: 33rd ACL (Student Session); 1995; 1995. p. 329-31.
35. Dalianis H. Concise natural language generation from formal specifications. PhD-Thesis. Stockholm: Dept Comput Systems Science-Royal Institute of Technology, Stockholm University; 1996 6/1996. Report No.: 96-008.
36. Dalianis H, Hovy E. On lexical aggregation and ordering. In: Sixth European Workshop on Natural Language Generation; 1997; Duisburg; 1997. p. 17-27.
37. Shaw J. Segregatory coordination and ellipsis in text generation. In: 36th Annual Meeting of the Association for Computational Linguistics and the 17th International Conference on Computational linguistics (COLING-ACL 98); 1998; 1998. p. 1220-6.
38. Reape M, Mellish C. Just what is aggregation anyway? In: 7th European Workshop on Natural Language Generation (EWNLG 99); 1999; Toulouse; 1999. p. 20-9.
39. Shaw J. Clause Aggregation using linguistic knowledge. In: 9th International Workshop on Natural Language Generation; 1998; Niagara-on-the-Lake, Canada; 1998.
40. Grosz B. The representation and use of focus in dialogue unterstanding. Technical Report. Menlo Park: SRI International; 1977. Report No.: 151.

41. Kamp H. A theory of truth and semantic interpretation. In: Truth, Interpretation and Information. Dordrecht: Foris Publications; 1981.
42. Grice HP. Logic and conversation. In: Speech acts. Cole P, Morgan JL, eds. New York, USA: Academic Press; 1975. p. 41-58.
43. Searle JR. Speech Acts: An essay in the philosophy of language. Cambridge: Cambridge University Press; 1969.
44. Austin JL. How to do things with words. Oxford: Clarendon Press; 1962.
45. Akmajian A, Demers RA, Farmer AK, Harnish RM. Linguistics, an introduction to language and communication. Cambridge, MA, USA: MIT Press; 1990.
46. Stent AJ. Aspects of natural language generation. Technical Report. Rochester: University of Rochester; 1998 11/1998. Report No.: 701.
47. Bateman J, Rondhuis KJ. Coherence relations: analysis and specification. project deliverable. Darmstatd: GMD-IPSI; 1994 6. 10. 1994. Report No.: R1.1.2: a, b.
48. Bateman J, Zock M. Updated list of generation systems; 2001. www.fb10.uni-bremen.de/anglistik/langpro/NLG-table/NLG-table-root.htm
49. Goldberg E, Driedger N, Kittredge R. Using natural language processing to produce weather forecasts. IEEE Expert 1994; 9 (2): 45-53.
50. Reiter E, Dale R. Building applied natural language generation systems. J Natural Language Engineering 1997; 3 (1): 57-87.
51. Paiva DS. A survey of applied natural language generation systems. Technical Report. Brighton: Information Technology Research Institute; 1998 7/1998. Report No.: ITRI-98-03.
52. Cahill L, Evans R, Mellish C, Reape M, Scott D. Towards a reference architecture for natural language generation systems – The RAGS project. Technical Report. Brighton: Inform Technology Res Institute; 2000 8/1999.
53. Elhadad M. The FUF Functional unifier: user's manual (version 3.0). Technical Report. New York: Columbia University, Dept. of Computer Science; 1990. Report No.: CUCS-012-91.
54. Elhadad M, Robin J. An overview of SURGE: a reusable comprehensive syntactic realization component. Technical Report. Beer Sheva, Israel: Dept. Comput Science, Ben Gurion University; 1996 4/1996. Report No.: 96-03.
55. Bateman J. Enabling technology for multilingual natural language generation: The KPML development environment. Natural Language Engineering 1997; 3: 15-55.
56. Lavoie B, Rambow O. a fast and portable realizer for text generation systems. In: Conference on Applied Natural Language Processing (ANLP'97); 1997 July 1997; Washington, DC: Melamed, I. D; 1997. p. 265-8.
57. CoGenTex. Tools; 2001. www.cogentex.com/technology/tools.html
58. Wilkinson J. Aggregation in natural language generation: Another look (M. Sc.). Waterloo: University of Waterloo, Canada; 1995.
59. Rösner D, Stede M. Generating multilingual documents from a knowledge base: The TECHDOC Project. In: 15th International Conference on Computational Linguistics (COLING 94); 1994; Kyoto; 1994.
60. Wagner JC, Rassinoux A-M, Baud RH, Scherrer J-R. Generating noun phrases from a medical knowledge representation. In: MIE 94; 1994; Lisbon, Portugal: Springer; 1994. p. 218-23.
61. Wagner JC, Richardson WD, Michel P-A, Baud RH, Rector AL, Scherrer J-R. Multilingual natural language generation as part of a medical terminology server. In: MEDINFO 95. Greenes RA, Peterson HE, Protti DJ, eds. Coach Publishing, Vancouver; 1995. p. 100-4.
62. Rector A, Zanstra PE, Solomon DW, Rogers JE, Baud R, Ceusters W, et al. Reconciling users needs and formal requirements: issues in developing a reusable ontology for medicine. IEEE Transactions Inform Biomed 1998; 2 (4): 239-42.
63. Dale R. Generating referring expressions: Constructing Descriptions in a domain of objects and processes. Cambridge, MA, USA: MIT Press; 1992.
64. Shaw J, McKeown K. Generating referring quantified expressions. In: 1st International Natural Language Generation Conference; 2000; Mitzpe Ramon, Israel; 2000. 100-7.
65. Eugenio Bd, Moore JD, Rosé CP. Generating 'distributed' referring expressions: an initial report. In: Eighth Int Workshop Natural Language Generation (INLG 96); 1996; Herstmonceux, France; 1996.
66. Deemter Kv. Generating vague descriptions. In: Int Natural Language Generation Conference (INLG-2000); 2000; Mitzpe-Ramon, Israel; 2000. p. 179-85.
67. Yeh C-L, Mellish C. An empirical study on the generation of anaphora in chinese. Computational Linguistics 1997;23(1):169-90.
68. Jameson A, Wahlster W. User modelling in anaphora generation: ellipsis and definite description. In: Europ Conference Artificial intelligence (ECAI 82); 1982; Orsay, France; 1982. p. 222-7.
69. Dale R, Reiter E. Computational interpretations of the Gricean Maxims in the generation of referring expressions. Cognitive Science 1995;19:233-63.
70. Dale R. Generating one-anaphoric expressions:Where does the decision lie? In: PACLING-II; 1995; Brisbane, Australia; 1995. p. 49-58.
71. Partee BH. Nominal and temporal anaphora. Linguistics Philosophy 1984;7:243-86.
72. Geldof S, Velde Wvd. An architecture for template based hypertext generation. In: 6th Europ Workshop Natural Language Generation (EWNLG 97). Hoeppner W, ed. Duisburg; 1997. p. 28-37.
73. DiMarco C, Hirst G, Hovy E. Generation by selection and repair as a method for adapting text for the individual reader. In: Workshop on flexible Hypertext (8th ACM International Hypertext Conference; 1997 May 1997; Southampton; 1997.
74. Wahlster W, Andre E, Finkler W, Profitlich H, Rist T. Plan-based integration of natural language and graphics generation. Artificial Intelligence 1993; 63: 387-427.
75. Andre E, Finkler W, Graf W, Rist T, Schauder A, Wahlster W. The automatic synthesis of multimodal presentations. In: Maybury M, ed. Intelligent Multimedia Interfaces. Menlo Park: MIT Press; 1993. p. 75-93.
76. Reiter E, Osman L. Tailored patient information: some issues and questions. In: From research to commercial applications: Making NLP technology work in practice (ACL-1997 workshop). Madrid; 1997. p. 29-34.
77. Hirst G, DiMarco C, Hovy E, Parsons K. Authoring and generating health-education documents that are tailored to the needs of the individual patient. In: 6th Int Conference User Modeling; 1997; Sardinia, Italy; 1997. p. 107-18.
78. Elhadad N, McKeown KR. Towards generating patient specific summaries of medical articles. 2001.
79. McKeown K, Pan S, Shaw J, Jordan D, Allen B. Language generation for multimedia healthcare briefings. In: Fifth ACL Conference on Applied Natural Language Processing (ANLP-97); 1997 31.3-4.3. Washington D.C.; 1997. p. 277-82.
80. Li P-Y, Evans M, Hier D. Generating medical case reports with the linguistic string parser. In: Fifth National Conference on Artificial Intelligence (AAAI 86); August 11-15, 1986; Philadelphia, PA: Morgan Kaufmann; 1986. p. 1069-73.
81. Giere W. BAIK - Befunddokumentation und Arztbriefschreibung im Krankenhaus. Taunusstein: Media; 1986.
82. McDonald CJ, Tierney WM, Blevins L. The benefits of automated medical record systems for ambulatory care. In: Implementing Health Care Inform Systems. Orthner HF, Blum BI, eds. New York: Springer; 1988.
83. Kraus D, Miche E. Arztbrief – generating medical reports in a multimedia environment. In: Medical Informatics Europe (MIE 91). Adlassnig KP, Grabner G, Bengtsson S, Hansen R, eds. 1991; Wien: Springer; 1991. p. 900-4.
84. Bernauer J, Gumrich K, Kutz S, Linder P, Pretschner P. An interactive report generator for bone scan studies. In: Proceedings Ann Symposium Computers Applied Medical Care (SCAMC 92). Clayton PD, ed. McGraw-Hill (IEEE Computer Society Press); 1992. p. 858-60.
85. Bernauer J. Conceptual graphs as an operational model for descriptive findings. In: SCAMC 92. Hanley & Belfus; 1992. p. 214-7.
86. Nowlan W, Kay S, Rector AS, Horan B, Wilson A. PEN&PAD: A multilingual patient care workstation based on a unified representation of the medical record and medical terminology. In: MIE 91; 1991; Vienna, Austria: Springer; 1991. p. 1043-8.
87. Bullock JC. Text generation from semantic network based medical records (Master's thesis). Manchester: University of Manchester; 1994.

88. Gregory J, Mattison JE, Linde C. Naming notes: transitions form free text to structured entry. Methods Inf Med 1995;34(1-2):57-67.
89. Birkmann C, Diedrich T, Ingenerf J, Rogers J, Moser W, Engelbrecht R. A formal model of diabetological terminology and its application for data entry. In: Med Inform Europe 97 (MIE '97). Pappas C, Maglaveras N, Scherrer J-R, eds. Amsterdam: IOS Press; 1997. p. 426-30.
90. Tange HJ, Hasman A, Robbé PFdV, Schouten HC. Medical narratives in electronic medical records. Int J Med Informatics 1997;46:7-29.
91. Reiter E, Dale R. Building natural language generation systems. Cambridge: Cambridge University Press; 2000.
92. Stede M. Lexical semantics and knowledge representation in multilingual text generation: Kluwer Academic Publishers; 1999.
93. Ceusters W, Spyns P, Moor GD, Martin W. Syntactic-semantic tagging of medical texts: The Multi-TALE Project. Amsterdam: IOS Press; 1998.
94. Coch J. Overview of AlethGen. In: 8th Int Workshop natural language generation (INLG 96); 1996; Herstmonceux, France; 1996. p. 25-8.
95. Calder J, Evans R, Mellish C, Reape M. "Free choice" and templates: How to get both at the same time. In: "May I speak freely?" Between templates and free choice in natural language generation. Becker T, Busemann S, eds. Saarbrücken: DFKI Saarbrücken report D-99-01; 1999. p. 19-24.
96. Bangalore S, Rambow O, Whittaker S. Evaluation metrics for generation. In: International Natural Language Generation Conference (INLG 2000); 2000; Mitzpe Ramon, Israel; 2000.
97. Rector AR. Clinical terminology: Why is it so hard? Methods Inf Med 1999; 38 (4-5): 239-52.
98. Mori AR. Archetypes of situations for clinical information - Terminology; 2001. www.intratext.com/y/CNR0001.htm
99. Rossi-Mori A, Consorti F, Galeazzi E. Standards to support development of terminological systems for healthcare telematics. In: Conference on Natural Language and Medical Concept Representation. WG6 I, ed. Methods Inf Med 1998; 37: 551-63.
100. Gangemi A, Pisanelli DM, Steve G. Some requirements and experiences in integrating terminological ontologies in medicine. In: KAW 98; 1998.
101. Haux R. Health care in the Information Society: What should be the role of medical informatics? Methods Inf Med 2002;1:31-4.
102. Lei Jvd. Closing the loop between clinical practice, research, and education. The potential of electronic patient records. ((■■■ dieser ursprüngliche Titel ist unstimmig: Medical informatics: challenges ■■■)). Methods Inf Med 2002;41:51-4.
103. O'Keefe QEW, Simborg DW, Epstein WV, Warger A. A computerized summary medical record system can provide more information than the standard medical record. ■■■ 1985; 254: 1185-92.
104. Kuhn K, Gaus W, Wechsler JG, Janowitz P, Tudyka J, Kratzer W, et al. Structured reporting of medical findings: evaluation of a system in gastroenterology. Methods Inf Med 1992;31:268-74.
105. Robertson R. The STOP computer system for generating tailored letters to help people stop smoking; 1999. www.csd.abdn.ac.uk/research/stop/ ComputerSystem.htm

Correspondence to:
Dr. Dirk Kraus
Dept. of Hospital Informatics
Heart and Diabtes Northrhine-Westphalia
Georgstrasse 11
D- 32545 Bad Oeynhausen
Germany
E-mail: dirk.kraus@suregen.de