

A Preliminary Analysis of the Dental Informatics Literature

T.K. Schleyer^{1*}, P. Corby², A.L. Gregg³

¹Center for Dental Informatics, School of Dental Medicine, University of Pittsburgh, 3501 Terrace Street, Pittsburgh, PA 15261; ²Division of Pediatric and Developmental Dental Sciences, School of Dental Medicine, University of Pittsburgh; ³Health Sciences Library System, Falk Library of the Health Sciences, University of Pittsburgh; *corresponding author, titus@pitt.edu

Adv Dent Res 17:16-20, December, 2003

Abstract — Dental informatics is an emerging discipline applying computer and information science to dental practice, research, education, and management. To date, the dental informatics research literature has not been comprehensively reviewed. This study reports an initial analysis of the dental informatics literature. We developed an initial, comprehensive retrieval strategy to locate dental informatics citations in MEDLINE (1966–April 2003), including three concepts: dentistry, computers, and research. After refinement of the search, we manually classified the final set into four categories: (1) non-dental; (2) dental, but neither dental informatics nor IT-related; (3) dental informatics; and (4) IT in dentistry. We analyzed informatics and IT-related citations regarding their distribution across journals, growth rate, the number of authors and their publication frequency, and content as expressed by Medical Subject Headings (MeSH). The final set of citations (n = 3872) consisted of: 12% non-dental articles; 59% dental, but not informatics- or IT-related articles; 16% informatics-related articles; and 13% IT-related articles. Informatics-related citations appeared in 176 journals, and IT-related citations in 206 journals. Approximately 50 papers are currently published in both categories yearly. While a great many authors have contributed to this literature, very few have published more than three papers. Main topics of articles included "Imaging and Image Processing", "Computer-aided Diagnosis and Therapy", "Computer-aided Instruction", and "Other". The dental informatics literature is small, but growing. Imaging and image processing predominate as research topics.

Introduction

Dental informatics is a new scientific discipline that is still in its formative stages (Schleyer and Spallek, 2001) yet shows signs of progressing maturation. Growing numbers of researchers are active in dental informatics research, venues for professional exchange and communication are becoming more numerous, and the quantity of informatics-related research papers is on the rise. Two NLM/NIDCR-funded dental informatics training programs continue to graduate a small but steady stream of qualified dental informatics researchers.

While the dental informatics literature is still sparse, the literature is an important resource in describing new scientific fields (Bush, 1996; Morris and McCain, 1998). The analysis of papers and articles can be used to define a field, determine its boundaries to other fields, describe the research problems and topics it addresses, and identify trends. An analysis of the biomedical informatics journal literature in 1998 (Morris and McCain, 1998) used intercitation and co-citation analyses to elucidate the field's structure. In examining citation patterns among 20 biomedical informatics journals, the authors found that the major focus areas in the field included biomedical engineering, biomedical computing, decision support, and education. The authors noted that none of the journals in

specific application areas of informatics, such as nursing and dentistry, ranked high enough in the exploratory analyses to be included in the final analysis. This reflects the fact that even until today no dedicated dental informatics journal (with the possible exception of the *Journal of Computerized Dentistry*, published by Quintessence Publishing, Inc.) exists.

However, the number of dental informatics papers in journals such as the *Journal of the American Medical Informatics Association*, the *Journal of the American Dental Association*, and the *Journal of Dental Education* is rising. We therefore concentrated on the single paper as the unit of analysis in our study. The scattering of the dental informatics literature among many journals resembles the dispersion of the biomaterials literature found in 1996 (Bush, 1996). Biomaterials scientists listed 65 journals that were of interest to their work. As a consequence, the retrieval of relevant papers was seen as a major obstacle in this emerging discipline.

We anticipated similar retrieval problems in the dental informatics literature. Since "dental informatics" first appeared in the MEDLINE-indexed literature only in 1986, it is intuitive that most work in this field is not yet labeled as such. As of this writing, searching for "dental informatics" in MEDLINE using the PubMed interface (National Library of Medicine, 1998) yielded about 200 citations. The same query for "medical informatics" retrieved over 295,000 citations. To the observer even only marginally familiar with both fields, the number of dental informatics citations appears artificially low. The retrieval problem for this literature is therefore evident at first glance.

The objective of this study was to perform an initial, general analysis of the dental informatics literature. Our first interest was how best to retrieve it. Second, we wanted to determine its size, the number of authors, its distribution across journals, its growth rate, and its general content as expressed by Medical Subject Headings (MeSH). We performed the same analyses with the information technology (IT) literature in dentistry, to gain an understanding about the relative characteristics of each of those literatures.

Methods

Two main considerations preceded the development of the retrieval and analysis protocol for this study. First, dental informatics is a research discipline. Similarly to biomedical informatics, it is frequently confused with information technology (Schleyer, 2003). Dental informatics is the application of computer and information science to improve dental practice, research, education, and management (Schleyer and Spallek, 2001) (biomedical informatics paper in these proceedings). (AQ) As a scientific field, it is concerned with model formulation, system design and development, system implementation, and evaluation (Friedman, 1995). Information technology, on the other hand, is primarily

Key Words

Dental informatics, information retrieval, literature, MEDLINE.

Presented at "Dental Informatics & Dental Research: Making the Connection", a conference held in, Bethesda, MD, USA, June 12-13, 2003, sponsored by the University of Pittsburgh Center for Dental Informatics and supported in part by award IRI3DE014611-01 from the National Institute of Dental and Craniofacial Research/National Library of Medicine.

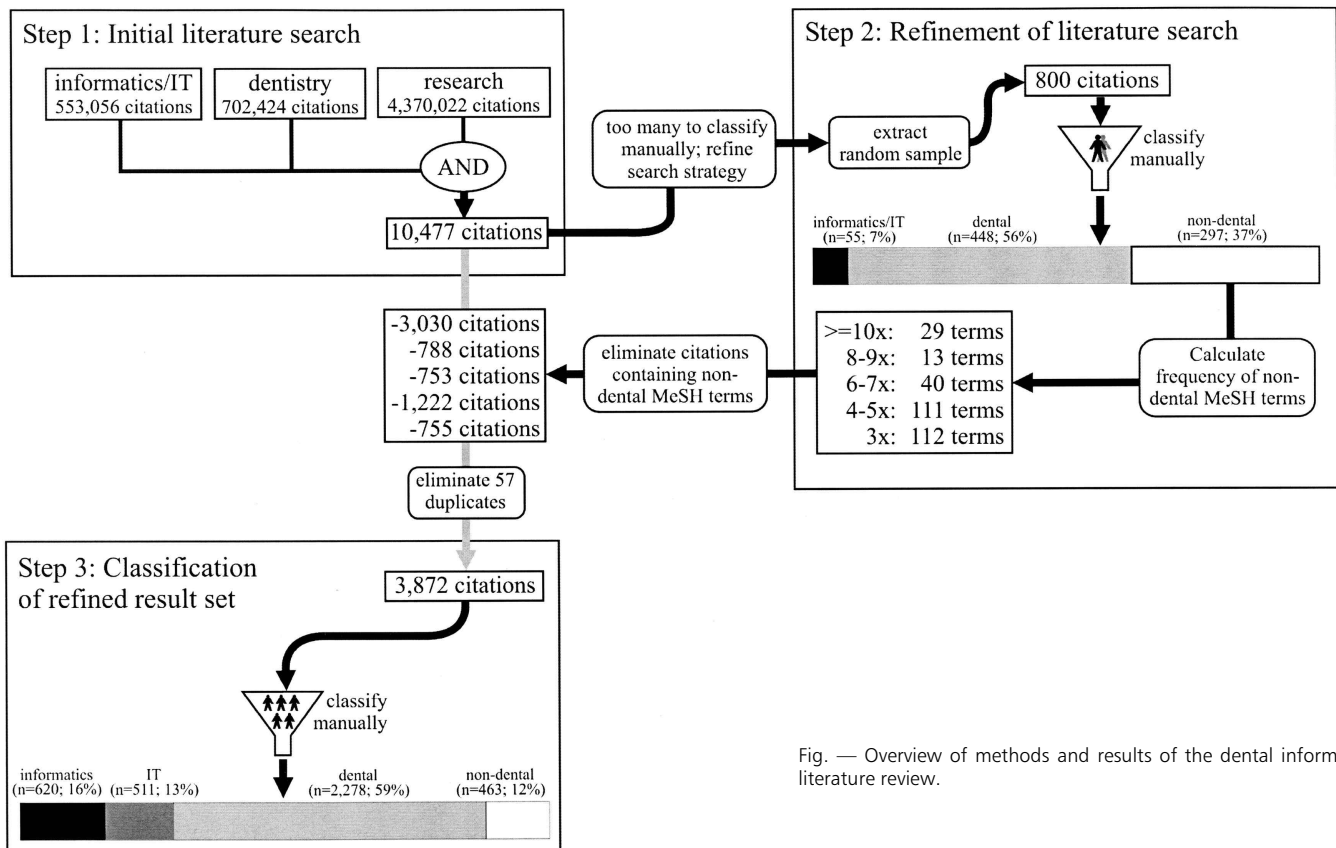


Fig. — Overview of methods and results of the dental informatics literature review.

focused on the implementation, application, and support of computer technology and telecommunications. While the two fields share some activities, the motivation for dental informatics is fundamentally different from that of information technology. It is the advancement of the state of the art through methods of scientific inquiry. Since developing an effective retrieval strategy for dental informatics alone was very difficult, we expected a certain number of articles about information technology in dentistry to appear in the result set. We decided to separate these two categories in the manual review of the final result set.

A second consideration is the relationship between dental informatics and its parent discipline, biomedical informatics (Schleyer, 2003). Dental informatics, just as other applied informatics fields, contributes to the basic theory and concepts underlying all of biomedical informatics. Generalizable methods become dissociated from the particular domain problem that may have given rise to their development, and thus become part of the methodological foundation of biomedical informatics. In examining a particular research paper, it might be difficult to label its subject "dental informatics" if it describes a universally applicable method. We therefore expected fundamental methods papers in our result set only if they referenced the dental research problem that resulted in the development of the new method.

Our initial search strategy for the dental informatics literature was comprised of three concepts: dentistry, computers, and research. This search strategy was intentionally broad, because high sensitivity (*i.e.*, not missing any potentially relevant citations) was very important. For dentistry, we included 48 high-level Medical Subject Headings (MeSH) terms, such as "stomatognathic system", "dental records", "dental research", and "education, dental". All high-level headings were exploded in the search. The concept "computers" included eight search terms, such as "technology", "medical informatics", "dental informatics", and "computers" truncated to capture all citations containing words that began

with "comput". The concept of "research" presented a special problem, since it is a meta-concept that is not normally applied to citations in MEDLINE (Bartling *et al.*, 2003). We attempted to cover this subset by searching for research-related terms (such as "statistics", "results", "study", and "methods") in titles and abstracts. We excluded the publication types "comment", "editorial", and "letter" from the search results. We also eliminated articles without abstracts, which biased the results toward more recent publications. Since abstracts were crucial to determining the content of publications, we accepted this limitation. We placed no restrictions on the language in which articles had been published.

We ran this initial search strategy on the MEDLINE database *via* Ovid from 1966 to April, 2003. We then randomly sampled 800 papers from this initial result set, and two authors (TKS and PC) classified them jointly for content (see Fig.). The content classifications were (1) non-dental, (2) dental but neither dental informatics nor IT-related, (3) dental informatics, and (4) IT in dentistry. We discussed disagreements and reconciled them directly. The purpose of this initial review was to determine any difficulties or problems in classifying papers, and to estimate the percentage of relevant citations in the initial search results.

Based on an analysis of Medical Subject Headings (MeSH) term frequencies in the initial result set, we refined our search strategy. We mainly focused on citations that had been classified as "non-dental", to reduce the number of irrelevant citations. Using a computer program written in Python (see www.python.org), we extracted each MeSH term and calculated its frequency within the set. Starting with the most frequently occurring terms, we examined each term regarding its possible elimination from the result set. This step allowed us to reduce the number of inappropriate citations in the set while still retaining all relevant citations. Thus we increased the specificity of our search with no change in sensitivity.

After eliminating all irrelevant terms, we re-ran the refined MEDLINE search, and examined the result set again

TABLE 1 — Number of Publications in Dental Informatics Research and IT in Dentistry per Journal Since 1975 (limited to journals with seven or more citations in dental informatics)

Journal	Dental Informatics	IT in Dentistry
<i>Dento-Maxillo-Facial Radiology</i>	56	24
<i>Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, & Endodontics</i>	69	21
<i>Journal of Dental Education</i>	25	11
<i>Journal of Clinical Periodontology</i>	22	8
<i>Journal of Prosthetic Dentistry</i>	18	14
<i>Caries Research</i>	15	2
<i>British Dental Journal</i>	11	9
<i>Journal of Dental Research</i>	11	9
<i>Acta Odontologica Scandinavica</i>	10	-
<i>Journal of Oral Rehabilitation</i>	10	2
<i>Journal of the American Dental Association</i>	10	9
<i>Journal of Periodontology</i>	9	11
<i>European Journal of Dental Education</i>	8	8
<i>Journal of Endodontics</i>	8	7
<i>Journal of Periodontal Research</i>	8	2
<i>IEEE Transactions on Biomedical Engineering</i>	7	2
<i>European Journal of Dental Education</i>	7	8(AQ)

TABLE 2 — Yearly Publication Output in Dental Informatics and IT in Dentistry, Including the Rate of Change in Percent (year to year) Since 1975

Year	Informatics	Change in %	IT	Change in %
1975	6	N/A	4	N/A
1976	6	0	3	-25
1977	5	-17	2	-33
1978	5	0	2	0
1979	9	80	1	-50
1980	6	-33	4	300
1981	9	50	6	50
1982	10	11	3	-50
1983	9	-10	4	33
1984	11	22	4	0
1985	18	64	8	100
1986	19	6	10	25
1987	16	-16	8	-20
1988	25	56	17	113
1989	26	4	33	94
1990	12	-54	26	-21
1991	15	25	24	-8
1992	29	93	24	0
1993	35	21	31	29
1994	28	-20	29	-6
1995	38	36	33	14
1996	30	-21	26	-21
1997	28	-7	20	-23
1998	33	18	26	30
1999	48	45	32	23
2000	59	23	39	22
2001	36	-39	54	38
2002	49	36	37	-31
Total	620		510	

regarding the proportion of relevant articles. (The final MEDLINE search strategy and result set for dental informatics and dental IT papers are available online at <http://di.dental.pitt.edu/DIlitreview/>).

At this point, it was impossible to refine the search strategy further without expending significant effort, and we decided to filter the final set using a small group of expert raters (dental informatics faculty and post-graduate students). Each rater was assigned several hundred citations, which they classified independently.

We then analyzed the resultant set of dental informatics papers by computing MeSH term frequencies, and calculating

journal counts and author frequencies. We also plotted the yearly publication output in both the informatics and IT categories. These analyses provided an initial summary view of both the dental informatics and dental IT literatures.

Results

The phases and results of the literature retrieval and classification are summarized in the Fig. The initial, broad search strategy (Fig., Step 1) incorporating the concepts "dentistry", "computers", and "research" resulted in the retrieval of 10,477 citations. The random sample of 800 citations extracted from this set (Fig., Step 2) consisted of 37% non-dental articles (297 citations), 56% dental, but not informatics-or IT-related, articles (448 citations), and 7% informatics-or IT-related articles (55 citations).

The analysis of MeSH term frequencies in the set of non-dental articles determined that 72 terms occurred more than ten times, 25 terms eight or nine times, 60 terms six or seven times, 153 terms four or five times, and 147 terms three times. From this total of 457 terms, we were able to eliminate 305 inappropriate terms. Excluding these terms from the search reduced the size of the result set by 6548 citations. Examples of eliminated terms included "Sleep Apnea Syndromes", "Orbit", "Hip Prosthesis", and "Pharyngitis". The final raw set consisted of 3929 citations.

After the elimination of 57 duplicates (Fig., Step 3), the final set contained 3872 citations. A group of two dental informatics faculty and three post-graduate students in dental informatics manually classified the citations in this set. Due to resource constraints, each citation was reviewed by only a single individual. The final set contained 12% non-dental articles (463 citations), 59% dental, but not informatics-or IT-related, articles (2278 citations), 16% informatics-related articles (620 citations), and 13% IT-related articles (511 citations). Compared with our initial result set, our filtering approach reduced the proportion of non-dental articles by more than two-thirds, left the proportion of dental articles about the same, and increased the proportion of informatics-and IT-related citations by a factor greater than four.

We then analyzed the set of 620 informatics-related citations and 511 IT-related citations as described above. As a consequence of retrieving only citations with abstracts, the earliest citation in both sets was from 1975. While research on computer applications in dentistry had been published prior to 1975, it could not be retrieved with our methods. The informatics-related citations appeared in 176 journals, and the IT-related citations in 206 journals. Journals with the most informatics-related citations (see Table 1; limited to journals with seven or more citations for dental informatics) included *Dento-Maxillo-Facial Radiology* (n = 56), *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, & Endodontics* (n = 69), and the *Journal of Dental Education* (n = 25). Journals with the most IT-related citations (limited to journals with seven or more citations) included *Dento-Maxillo-Facial Radiology* (n = 24), *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, & Endodontics* (n = 21), and the *Journal of Prosthetic Dentistry* (n = 14).

Table 2 shows the yearly publication output in the categories of interest. Currently, approximately 50 papers are published annually in each of the categories dental informatics and IT. Between 1975 and 2002, the historical yearly growth rate for informatics-related papers has been approximately 19%, while it was 20% for IT-related papers. However, mainly due to the small numbers of publications, earlier years show a much higher variability in the growth rate than later years. For the last ten years, the yearly growth rate for dental informatics was approximately 10%, and for IT 11%.

The number of unique authors in dental informatics was 1672, and the number of unique authors in IT in dentistry was 1349. Forty-two authors, or 2.5%, published four or more papers in dental informatics (accounting for 282 papers, or 17% of the total), while 27 authors, or 2%, published four or more papers in IT in dentistry (accounting for 122 papers, or 9% of the total).

The major MeSH terms according to which dental informatics papers are indexed fall into roughly four categories (see Table 3). "Imaging and Image Processing" is a major topic in 129 (or 72%) of the papers. Primarily, these papers are concerned with digital dental radiography. However, other imaging modalities—such as computed tomography, magnetic resonance imaging, and three-dimensional imaging—are also represented. With 22 occurrences (or 12%), "Computer-aided Diagnosis and Therapy" is the second most frequently treated topic. "Computer-aided Instruction", with eight occurrences (or 4%), is the third most frequent category. Twenty-one terms (or 12%) were too generic to be included in any other category.

Major MeSH terms provide only a very general description of the content of a MEDLINE citation. Typically, papers are indexed with between one and three major MeSH terms, and between five and 15 additional MeSH terms. The personnel indexing the papers sometimes must generalize significantly when assigning the major MeSH terms, and thus these terms may be only an imprecise reflection of the actual content. When we analyzed the frequencies for the non-major MeSH terms, their distribution reflected the same general categories as that of the major MeSH terms. We do not present these results here, however, since other analyses, such as clustering, would provide a more meaningful representation of the literature content.

Discussion

This study has shown that the dental informatics literature is small but growing. The total size of the dental informatics literature since 1975 is approximately 620 research papers. For reasons discussed below, this number probably underestimates the true total to some degree. The current yearly publication output in dental informatics is approximately 50 papers. During the past ten years, the number of papers has grown by 10% annually. Compared with other dental disciplines, the total publication output of informatics is quite small. A recent study (Yang *et al.*, 2001) determined that the publication output from 1989 to 1998 in seven dental disciplines ranged from 327 papers (endodontics) to 2765 papers (oral medicine) *per year*. The publication growth rate of all seven disciplines was 3% *per year*. Thus, it appears that dental informatics is a relatively fast-growing segment of the dental literature.

Dental informatics papers are spread across a large number of journals, and this makes it difficult to define a cohesive literature. Some journals, such as the *Journal of the American Dental Association*, have established a section for informatics, which may serve as a better focal point for informatics-related papers. The number of very active authors, both currently and historically, is quite low. The number of authors with fewer than three total publications is very high, and may reflect the broad but relatively superficial interest in this discipline. At this point, it seems premature to assume that the dental informatics community could support one or more journals of its own. Not enough trained dental informaticians exist to produce a steady stream of research papers of high quality. In addition, it appears more useful to introduce dental informatics research to a mainstream audience in dentistry than to provide an additional communication forum for a very small group of individuals.

The predominant area of research in dental informatics is imaging and image processing. This is intuitive, inasmuch as

TABLE 3 — Major MeSH Terms in Dental Informatics Papers and Their Frequencies Grouped into the Categories "Imaging and Image Processing", "Computer-aided Diagnosis and Therapy", "Computer-aided Instruction", and "Generic Terms"

Topic	# of Occurrences
Imaging and Image Processing	129 (72%)
Radiographic image enhancement	37
Image processing, computer-assisted	32
Tomography, x-ray computed	28
Radiography, dental, digital	21
Imaging, three-dimensional	7
Magnetic resonance imaging	4
Computer-aided Diagnosis and Therapy	22 (12%)
Radiographic image interpretation, computer-assisted	6
Diagnosis, computer-assisted	5
Therapy, computer-assisted	5
Image interpretation, computer-assisted	2
Computer-aided design	2
Expert systems	1
Surgery, computer-assisted	1
Computer-aided Instruction	8 (4%)
Computer-assisted instruction	4
Education, distance	2
Educational technology	2
Generic Terms	21 (12%)
Information systems	2
Automatic data processing	2
Internet	2
Computer graphics	2
Electronic mail	2
Computer systems	1
Software	1
Computer communication networks	1
Hospital information systems	1
Information management	1
Information services	1
Information storage and retrieval	1
Printing	1
Radiology information systems	1
Video recording	1
Videodisc recording	1

imaging, in its various forms, plays a significant role in clinical dental care. Among the major subject areas, computer-aided diagnosis and therapy and computer-aided instruction also receive significant attention in the literature. Interestingly, the top-level classification of research topics in dental informatics closely resembles that found in biomedical informatics (Morris and McCain, 1998), despite the fact that those classifications were determined by very different methods. A more fine-grained analysis of research topics in dental informatics will be the subject of a future study.

This study has several important limitations. The total number of dental informatics citations worldwide since 1975 is most likely higher than 620. Several circumstances are responsible for this limitation. In our literature search, we focused on the MEDLINE database only. Other databases, such as the *Science Citation Index*, and hand-searching non-indexed journals would have provided additional citations. The number of publications on IT in dentistry is probably underestimated to a greater extent than the number of publications in dental informatics. The reason for this circumstance is that our search focused on research publications, and thus eliminated many publications on IT that did not match our search criteria. However, since informatics was our primary focus, this limitation is acceptable.

It is uncertain whether research in countries other than the US is represented to the same degree as research published in

the US. Aside from the US, Europe is a major contributor to the informatics literature. Another limitation was the fact that we excluded citations without abstracts. This decision eliminated from our result set the early work in dental informatics dating back to the 1960s and early 1970s. Finally, we do not have a true measure for the recall (or sensitivity) of our search strategy. The MEDLINE database may contain additional relevant dental informatics citations that we missed by our choice of search terms. We attempted to counteract this possible limitation by formulating our initial search strategy as broadly as possible, and then eliminating only irrelevant citations from the results.

As a follow-up to this study, we plan to perform a more detailed review and classification of the dental informatics literature. For instance, one area of interest is where dental informatics has made specific contributions to the theoretical biomedical informatics literature. Another goal of this subsequent study would be to examine in more detail the research that has been performed in dental informatics, its results, and the implications for future research.

Acknowledgments

The authors thank William Bartling, Heiko Spallek, and Humberto Torres-Urquidy for their assistance in classifying papers in the final result set. In addition, the authors thank

Rishi Bhatnagar, Thomas Carnaggio, Heiko Spallek, and Astrid Weirauch for their assistance in developing the Figure.

References (AQ)

- Bartling WC, Schleyer TK, Visweswaran S (2003). Retrieval and classification of dental research articles. *Adv Dent Res* 17:115-120.
- Bush RB (1996). Biomaterials: an introduction for librarians. *Sci Technol Libr* 15(4):3-28.
- Friedman CP (1995). Where's the science in medical informatics? *J Am Med Inform Assoc* 2:65-67.
- Morris TA, McCain KW (1998). The structure of medical informatics journal literature. *J Am Med Inform Assoc* 5:448-466.
- National Library of Medicine (1998). Welcome to PubMed. [Online]. Available: <http://www.ncbi.nlm.nih.gov/PubMed/>. Accessed 4-24-1998.
- Schleyer TK (2003). Dental informatics: an emerging biomedical informatics discipline. *J Dent Educ* 67:1193-1200.
- Schleyer T, Spallek H (2001). Dental informatics. A cornerstone of dental practice. *J Am Dent Assoc* 132:605-613.
- Yang S, Needleman H, Niederman R (2001). A bibliometric analysis of the pediatric dental literature in MEDLINE. *Pediatr Dent* 23:415-418.