J Med Syst (2006) 30:153–157
DOI 10.1007/s10916-005-7993-0

RESEARCH PAPER

Characterizing Cancer Information Systems

Roy Rada

Received: 25 April 2005 / Accepted: 22 June 2005

Abstract The objective is to determine the extent to which information systems (IS) for cancer are unique and necessary. Via an analysis of Medical Subject Headings used to index relevant literature and other bibliometric techniques, cancer IS are compared and contrasted with IS of other specialties. Cancer IS are relatively little discussed and primarily connect radiation equipment with the radiation oncology staff. By contrast, clinical laboratory and radiology IS are frequently discussed and connect specialized equipment to the hospital. A "Specialty Need" model accounts for these patterns and says that the "need for a specialty IS" is proportional to the "uniqueness of the specialty tools" plus the "degree to which the information from those tools is needed throughout the particular health care entity."

Keywords Cancer information systems · Bibliometrics · Radiation oncology · Management information systems

Introduction

Health care informatics can be described in diverse ways [1]. One way is to look at various information technologies, such as databases and networks [2]. Another way considers roles, such as physicians and administrators [3]. Functions of health care information systems (IS) include [4].

- administration, like accounting and scheduling,
- patient management, like medical record maintenance and order entry,
- general clinic support, such as laboratory or radiology support, and
- specialty clinical support, such as oncology and cardiology support.

Administration, patient management, and general clinic support functions each correspond with well-recognized IS. Under what conditions should a clinical specialty be supported by its own unique IS?

One of the earliest and best known oncology or cancer information systems (CIS) was developed in the 1970s [5]. The CIS focused on cancer chemotherapy. As the system aged, the question arose as to whether to continue to maintain it or to switch to using the hospital IS [6]. This practical problem leads to a generic question:

Does a cancer center need a CIS?

If one were to replace "cancer" with "radiology" or "pathology," then the answer would be yes [7]. However, if one replaces "cancer" with "cardiology," then the answer is less clear. The question naturally rests in a complex setting. For instance, cancer centers vary widely, from stand-alone, giant cancer centers, to small clinics that are part of another organization, and the software needs may vary.

This study focuses on the characteristics of CIS and tries to establish through bibliometric means the extent to which such systems are unique. Bibliometrics is the application of mathematics and statistical methods to books and other media of communication [8]. Scientometrics is the study of the measurement of scientific and technological progress and typically uses bibliometrics [9].

Bibliometric analysis

The bibliometric method uses empiric data from and quantitative analysis of published literature [10]. Present-day bibliometric research is either [11]

R. Rada (✉)
Department Information Systems, University of Maryland, Baltimore County,
Baltimore, MD 21250
e-mail: rada@umbc.edu

- basic research into bibliometric methods,
- bibliometrics for scientific discovery in a discipline, or
- bibliometrics for management in which structures of disciplines are assessed.

This study contributes to scientific understanding of the discipline of medical informatics and management of health care information systems. This study does not contribute to bibliometric methodology.

First-order patterns

The bibliometric database in this experiment is MEDLINE from the National Library of Medicine. MEDLINE contains citations for millions of journal articles. Each citation is associated with a set of Medical Subject Heading (MeSH) terms that describe the content of the associated article. MeSH is a thesaurus used for indexing articles and for searching MeSH-indexed databases, in particular, MEDLINE [12].

Indexers for MEDLINE typically choose about a dozen MeSH terms to represent a journal article [13]. About two of these will be indicated as the principal focus of the article by appending the attribute of "Major" to the term in the indexing. A MEDLINE search can specify that articles are returned for which a particular MeSH concept was a "Major" concept in the article.

In the MeSH thesaurus hierarchy, the highest-level concept related to IS is "Information Science." One descendant of "Information Science" is "Management IS" which, in turn, has nine children that include "Clinical Laboratory IS" and "Radiology IS" (for the purposes of this presentation, "Clinical Laboratory IS" will be considered synonymous with "Pathology IS"). Neither "cancer IS" nor "cardiology IS" equates to exactly one MeSH term.

A query for CIS might combine the concepts "management IS" with MeSH terms that contain the word "oncology." In MeSH, for better or worse, the concept "radiation oncology" is a child of "medical oncology" and the only other term containing the term "oncology" is "hospital oncology service." A query with a concept will automatically expand to include all concepts that are descendants of the concept used in the query. Thus one query to represent "CIS" is

("medical oncology"[majr] OR "hospital oncology service"[majr]) AND "management information systems"[majr]

where a "major" term is indicated by "[majr]". This query when applied to MEDLINE in July 2004 retrieved 40 citations. By contrast the retrieval on 'radiology IS' retrieved 3,097 citations, while the query on "clinical laboratory IS" retrieved 1,286 citations.

The term cardiology occurs in only the two MeSH terms "cardiology" and "hospital cardiology service." A query with those two terms combined with "management information systems" was used to capture "cardiology IS" articles. This query retrieved 41 citations in July 2004—essentially the same number as for CIS.

Co-occurring patterns

A common method in bibliometrics is to look for co-occurrences. A potential relationship between two or more things depends principally upon how many common relationships people recognize between these things [14]. In the medical field, the co-occurrence of MeSH terms has been used to help recognize relationships among co-occurring concepts [15].

In this project, citations retrieved from MEDLINE were analyzed for co-occurring MeSH terms. A computer program was written in Active Server Pages that extracts information from citations. For each article, each "major" MeSH term was stored in a table along with a unique identifier for the article. The program determined how many times each pair of MeSH terms occurred in an article's indexing. Finally, the program displayed a list of the co-occurring MeSH term pairs, in descending order of the number of times they co-occurred.

Indexers also apply "qualifiers" to index terms. A qualifier serves to further characterize a concept. The program also accounted for co-occurring patterns of qualifiers.

The indexing of the articles returned for the CIS query showed that

- "medical oncology" co-occurred frequently with "database management system" and "medical record system,"
- "radiation oncology" co-occurred most with "radiology IS," and
- "hospital oncology service" co-occurred most with "clinical pharmacy" and "ambulatory IS."

The most frequently appearing qualifier was "organization & administration." These patterns make sense in that:

- medical oncology deals frequently with chemotherapy that involves the medical record and other databases,
- radiation oncology is closely related to radiology, and
- hospital oncology services have patients going to and from ambulatory clinics and needing pharmacy services.

These different relationships concern the organization and administration of the oncology services (as reflected in the most frequently occurring qualifier).

In the indexing of articles returned for the query "cardiology IS" the most frequently co-occurring MeSH major terms were "cardiology" and

- "computer communication networks,"
- "radiology IS,"
- "computer-assisted signal processing," and
- "computer-assisted image processing."

Cardiology IS are different from oncology IS in that they have a narrower focus on signals or images and networking. The most frequently appearing qualifier was "standards" which makes sense in light of the importance of standards in "communication networks" for cardiology signals or images.

In the citations retrieved to the query "radiology information systems[majr]" the most frequently co-occurring MeSH terms with "radiology IS" involved:

- connecting to the hospital IS and
- communicating images.

With the query "clinical laboratory information systems [majr]" the terms most frequently co-occurring with "clinical laboratory IS" involved communicating between the clinical laboratory and medical records. Thus, for "radiology IS" and "pathology IS" a connection to the hospital is more frequently addressed than for CIS or "cardiology IS."

Divergent query results

The "CIS" query addressed oncology applications of management information systems. How will the study results change when the query diverges and includes "neoplasm" and non-management aspects of information systems? The "IS" component of the CIS query was augmented with the children of "Information Science" (not just the children of "management IS"). The "oncology" component was augmented with the term "neoplasms" (and thus also all the descendants of "neoplasms"). After running the co-occurrence program on the citations retrieved by this broadened query, the most frequently co-occurring MeSH term pairs were related to "mass screening" and "breast neoplasm."

The query for "Cardiology IS" was also broadened like the "Cancer IS" query was, but now "neoplasms[majr]" was replaced with "cardiovascular diseases[majr] OR cardiovascular system[majr]." The retrieved citations emphasized "Cardiovascular Models" and "Computer-Assisted Signal Processing," as they did before the query was broadened.

The author experimented by narrowing the divergent queries. The intent was to further understand the intrinsic character of CIS by removing some of the query terms. However, the results are most meaningful with the queries already presented in the sections titled "first-order patterns" and "co-occurring patterns."

Discussion

While the radiology and pathology IS literature focuses on connectivity between the department and the rest of the health system, the oncology and cardiology IS literature focuses on diagnostic or treatment tools and methods specific to that clinical specialty. The extent to which a clinical specialty would need specialty-specific software would be partly related to the extent to which the unique diagnostic and treatment methods and tools of that specialty require unique software support. When that uniqueness exists and the information from those tools and methods furthermore needs to interface to the information infrastructures of the health system, such as the patient medical record system, then the commercial need for a unique specialty IS increases.

The preceding observations suggest the following 'Specialty Need' model:

> The "need for a unique, specialty IS" is proportional to the "uniqueness of the specialty tools and methods" plus the "degree to which the information from those tools and methods should be integrated with the information infrastructure of the health care entity."

Radiology and pathology have unique tools, and the rest of the health system needs to get orders to those departments and get results from those departments. For oncology and cardiology the need to integrate unique tools and methods with the rest of the health enterprise is less apparent.

Other perspectives

The conclusion from the bibliometric analysis has led to a tantalizing model about specialty software. Other sources of information will be mined to extend the answer to the question of whether CIS are unique. Next, content analysis of vendor web sites and provider interviews are presented.

Vendors

One approach to characterizing CIS is to study the vendors of CIS. Finding a list of vendors of CIS is not straightforward. One list was published in the "Health Informatics Resource Guide 2004 [16]." That Guide provides lists of vendors under different categories, such as radiology IS and oncology IS. The publication does not make clear how companies are entered into the lists. The author took the oncology, cardiology, radiology, and pathology lists of vendors and visited the web sites of each vendor. From the web sites, the author determined the extent to which the company had a product that met the needs of the specialty in question.

The oncology list has dozens of companies on it. Only a handful specified an oncology function and most of those were for radiation oncology. The lists entitled "radiology IS" and "pathology IS" contained a higher proportion of vendors offering IS specific to the list topic than either the oncology or cardiology lists did.

To go beyond the "Resource Guide," 16 students were each asked to find on the web five IS vendors that stress in their documentation that they support oncology. The students identified 28 different companies that marketed CIS, and of those CIS:

- 2 supported only chemotherapy,
- 15 supported only radiation, and
- 11 supported both radiation and chemotherapy.

The history of the largest vendors shows that they first supplied radiation equipment, then radiation oncology software, and finally medical oncology software. However, the diffusion of medical oncology software in the marketplace is not as extensive as radiation oncology software [17]." The cancer chemotherapy part of medical oncology software remains partially experimental.

One difficulty in diffusing medical oncology software is that it should communicate with the medical record, the pharmacy, the clinical laboratory, and scheduling. However, standards for supporting such communication are not widely accepted. Thus, installing medical oncology software is a large undertaking in a particular institution.

Most vendors do not offer CIS. Those vendors that address oncology are primarily addressing the radiation oncology business. The market of stand-alone cancer centers buying complete CIS has not been large enough to drive many vendors to that market, and for cancer centers within a larger entity, integrating a CIS with the larger entity is currently difficult.

The role of government research, health plans, and patients are next briefly addressed:

- The American National Cancer Institute (NCI) has initiated a national CIS project to facilitate access to molecular and clinical trials data [18].
- In the United States, "disease management [19]." is seen as part of the "health plan" enterprise, and while health plans generally do not have cancer-specific software, one disease management company focuses exclusively on cancer and has a CIS [20].
- No patients are more concerned about their diagnosis and treatment than are cancer patients, and many cancer patients have sought help from one another through online discussion groups [21].

These observations are consistent with the "Specialty Need" model and also show opportunities to extend CIS in the direction of research and consumer systems.

Providers

This author interviewed the major stakeholders in the CIS at the Johns Hopkins Cancer Center. That particular center is part of a large hospital complex, had developed its own CIS a quarter century earlier, and was spending $2 million per year on its internal, CIS staff for software maintenance [6]. The interviewees were asked whether their cancer center needed a unique CIS. Salient replies included:

- The medical creator of the IS said "No." While trend data is vitally important in cancer, trend data is needed in many specialties.
- The director of programming at the Center said "No." She said that while the cancer center needs are complex, they are not unlike the complex needs of other parts of the hospital dealing with chronically ill patients.
- The technical creator of the IS said "No" for the same reason as the director of programming.

Next, the author interviewed several people from stand-alone cancer centers, such as Memorial Sloan-Kettering. They felt that the needs of oncology, particularly for chemotherapy, required cancer-specific software. The uniqueness of cancer chemotherapy software is described in the literature [22, 23], but the literature also shows that stand-alone, cancer centers are most likely to have a CIS. The two largest stand-alone cancer centers in the US both invest large sums in acquiring or creating software that is tailored to their cancer-specific needs:

- Memorial-Sloan Kettering relies on tailored products from many different vendors, while
- M.D. Anderson Cancer Center relies extensively on home-grown software.

Even small, independent cancer centers want cancer-specific software [24]. Of course, in the final analysis what any provider wants is software that optimally supports the particular situation of that provider [25].

The "Specialty Need" model can be interpreted in a manner consistent with these interview results. That model says that a CIS is needed when the information from the cancer tools and methods should be integrated with the information infrastructure of the health care entity. The bibliometric data suggest that generally this integration of cancer data are not a top priority in a typical entity. However, in stand-alone cancer centers, integration of cancer data with the applications of the center is a top priority.

Conclusion

MeSH does not include a concept for CIS. However, MeSH does have single terms representing each of several other types of IS that might be comparable to CIS, such as "radiology IS" and "clinical laboratory IS." To investigate the attributes of CIS, MEDLINE queries were generated, and the MeSH indexing of the retrieved literature was analyzed.

The retrieved citations for CIS highlight the distinct concerns of medical oncology versus radiation oncology. By contrast, for "cardiology IS" the query retrieves citations that emphasize cardiovascular signal or image processing. Literature about radiology and clinical laboratory IS emphasizes connectivity within and among departments.

The interpretation of the bibliometric data leads to a "Specialty Need" model which states that a specialty needs unique software to the extent that its tools require software and the data from the tools needs to be shared across the health care entity. The need for a CIS would grow as an oncology department within a hospital provided radiation or chemotherapy treatments that entailed further communication with the hospital medical record. However, integrating a CIS with the hospital medical record is typically so costly that the benefits of cancer-specific software are often outweighed by the costs of integration. As standards for the medical record become more prevalent and make integration easier, specialty-specific software might become increasingly common.

While many oncology departments rely on hospital-wide IS, some stand-alone cancer centers have CIS built specifically for cancer. However, historically these cancer centers do not share a common software platform. The American National Cancer Institute has been funding large projects to develop shared software across cancer centers.

Acknowledgment Rona Abraham wrote the computer program, which determines co-occurrence of MeSH terms in MEDLINE citations.

References

1. Talmon, J., and Hasman, A., Medical informatics as a discipline at the beginning of the 21st century. *Methods Inf. Med.* 41(1):4–7, 2002.
2. Hatcher, M., and Heetebry, I., Information, technology in the future of health care. *J. Med. Syst.* 28(6):673–688, 2004.
3. Menachemi, N., Burke, D. E., and Ayers, D. J., Factors affecting the adoption of telemedicine: A multiple adopter perspective. *J. Med. Syst.* 28(6):617–632, 2004.
4. Rada, R., *Information Systems for Health Care Enterprises*, 2nd Edition, Hypermedia Solutions Limited, Baltimore, 2003.
5. Enterline, J., Lenhard, R., and Blum, B., (eds.), *A Clinical Information System for Oncology*, Springer–Verlag, New York, 1989.
6. Rada, R., and Finley, S., The aging of a clinical information system. *Biomed. Inform.* 37(4):319–324, 2004.
7. Pilling, J. R., Lessons learned from a whole hospital PACS installation. *Clin. Radiol.* 57(9):784–788, 2002.
8. Pritchard, A., Statistical bibliography or bibliometrics. *J. Doc.* 25(4):348–349, 1969.
9. Garfield, E., Scientometrics Comes of Age. *Curr. Contents* 46:5–10, 1979.
10. Estabrooks, C., Winther, C., and Derksen, L., Mapping the field: A bibliometric analysis of the research utilization literature in nursing. *Nurs. Res.* 53(5):293–303, 2004.
11. Glanzel, W., and Schubert, A., A new classification scheme of science fields and subfields designed for scientometric evaluation purposes. *Scientometrics* 56(3):357–367, 2003.
12. Backus, J., Davidson, S., and Rada, R., Searching for patterns in the MeSH vocabulary. *Bull. Med. Libr. Assoc.* 75(3):221–227, 1987.
13. Rada, R., Backus, J., Giampa, T., *et al*. Computerized guides to journal selection. *Inf. Technol. Libr.* 6(3):173–184, 1987.
14. Wren, J., Bekeredjian, R., Stewart, J., *et al*. Knowledge discovery by automated identification and ranking of implicit relationships. *Bioinformatics* 20(3):389–398, 2004.
15. Hristovski, D., Peterlin, B., Mitchell, J., *et al*. Using literature-based discovery to identify disease candida bte genes. *Int. J. Med. Inform.* 74(2–4):289–298, 2004.
16. Marietti, C., Resource guide. *Healthcare Inform.* 2003:1–108, 2004.
17. Joch, A., New software targets cancer therapy. *Healthcare Inform.* 2003:78–80.
18. Hartel, F., and de Coronado, S., Information standards within the National Cancer Institute. In: Silva, J., Ball, M., Chute, C., *et al* (eds.), *Cancer Informatics: Essential Technologies for Clinical Trials*, Springer, New York, 2002, pp. 135–156.
19. Zitter, M., A new paradigm in health care delivery: Disease management. In: Todd, W., and Nash, D. (eds.), *Disease Management: A Systems Approach to Improving Patient Outcomes*, American Hospital Association, Chicago, 1997, pp. 1–26.
20. Lee, F. C., Is cancer meant to be managed? *Manag. Care Cancer* 3(2):28–33, 2001.
21. Goldberg, H., and Safran, C., Support for the cancer patient: An internet model. In: Silva, J., Ball, M., Chute, C., *et al* (eds.), *Cancer Informatics: Essential Technologies for Clinical Trials*, Springer, New York, 2002, pp. 280–293.
22. Garde, S., Baumgarten, B., Basu, O., *et al*. A meta-model of chemotherapy planning in the multi-hospital/multi-trial-center-environment of pediatric oncology. *Methods Inf. Med.* 43(2):171–183, 2004.
23. Langlotz, C., Fagan, L., Tu, S., *et al*. A therapy planning architecture that combines decision theory and artificial intelligence techniques. *Comput. Biomed. Res.* 20(3):279–303, 1987.
24. Briggs, B., I.T. Helps Battle 'The Big C'. *Health Data Manag.* 4(5):1–4, 2004.
25. Wears, R. L., and Berg, M., Computer technology and clinical work: Still waiting for godot. *JAMA* 293(10):1261–1263, 2005.