The role of patient records in research: A bibliometric analysis of publications from an academic medical center in Taiwan

Hui-Chun Li[a], Yu-Wen Chen[b], Tzeng-Ji Chen[a,c,d,*], Shih-Hwa Chiou[c,e,f], Shinn-Jang Hwang[a,c]

[a]Department of Family Medicine, Taipei Veterans General Hospital, Taipei, Taiwan, ROC; [b]Medical Library, Department of Medical Education, Taipei Veterans General Hospital, Taipei, Taiwan, ROC; [c]School of Medicine, National Yang Ming Chiao Tung University, Taipei, Taiwan, ROC; [d]Big Data Center, Department of Medical Research, Taipei Veterans General Hospital, Taipei, Taiwan, ROC; [e]Institute of Pharmacology, National Yang Ming Chiao Tung University, Taipei, Taiwan, ROC; [f]Department of Medical Research, Taipei Veterans General Hospital, Taipei, Taiwan, ROC

Abstract

Background: This study aimed to assess the use of medical record items in clinical research in one large academic medical center in Taiwan.

Methods: A descriptive survey design was adopted to collect the data. Articles published in 2018 by Taipei Veterans General Hospital (TVGH) staff as the first author were obtained. The types of specialties and types of research were analyzed. To understand the conditions for the use of medical records, the retrospective research using hospital's medical records were analyzed. Each article was read in entirety to realize the use and number of patients and the medical record items.

Results: Among the 362 articles first-authored by TVGH staff in 2018, 219 (60.4%) were classified as clinical studies, 60 (16.6%) as basic studies, 53 (14.6%) as database studies, and 30 (8.2%) as other categories. About 50% of the retrospective research using TVGH medical records had a case number <100 (67 cases, 49.6%) with an average number of 41 cases and 13 studies (9.6%) had a case number >1000. Analysis of the number of medical record items used in 135 retrospective research studies based on TVGH medical records showed that 118 (87.4%) used basic patient information. In addition to basic information, notes written by professionals were used most frequently (73 cases, 54.0%), whereas medication information was used in 50 cases (37.0%); laboratory test data were used in 49 cases (36.2%); and body measurements was used in 27 cases (20%).

Conclusion: More than one-third of publications utilized medical records, but the patient numbers and record items in use were relatively limited. In the era of digitalization and big data analytics, the potential of medical records in research deserves attention. Investment in establishing a more accessible database of medical records to access nonstructural, descriptive medical records could be considered.

Keywords: Bibliometrics; Medical records; Publications

1. INTRODUCTION

Medical research is the foundation of medical science and can lead to a better understanding of disease, improved treatment outcomes, and a better patient prognosis.[1,2] Decades ago, researchers first observed and recorded the disease progression and treatment outcome for a single patient as a case report to learn more about the disease.[3] Since then, case series, cohort studies, randomized controlled studies, and other research and documentation methods have gradually developed. In addition to clinical research using retrospective chart analysis and prospective intervention, basic research, database research, and public health community research are also flourishing. Today, many studies no longer rely on the use of observation of the individual patient and documentation of the patient's medical records. However, medical records still have their role. These records document substantial information about the patient, including basic characteristics, disease progression, laboratory examinations, medications, treatments and surgeries, and past real-life experiences that can be directly referenced and applied.[4] In addition, medical records are legal documents that must be kept for at least 7 years, according to the Medical Care Act in Taiwan,[5] making it easier to access research data. The ability to collate and analyze medical records directly saves significant time and effort for researchers working to produce an article. As a result, medical research based on medical records still has a place in today's medical research environment.

This study was conducted at a large academic medical center in Taiwan. To determine the type of research conducted at the medical center and the status of the use of medical records in this research, this study analyzed the research articles published during 1 year (in 2018). The results of this study will contribute to the planning and

construction of a large database of medical records, known as the *Big Data Warehouse*.

2. METHODS

Established in 1958, Taipei Veterans General Hospital (TVGH) is one of the largest public medical centers in Taiwan, with approximately 6700 employees and 2.5 million medical visits per year.[6] Every year, TVGH produces a large number of research articles that are published in various journals and specialty publications. The medical library of the hospital routinely compiles the list of publications authored by the hospital staff. Especially for research articles published in the journals that are indexed in Science Citation Index (SCI) and Social Science Citation Index (SSCI), the librarians constructed a specific database with detailed items for systematic summarization.

The first step of this study was to obtain all staff publications from the library for the year of 2018. From this collection, all the research articles written by TVGH staff as the first author were identified. Articles for which the TVGH staff was the corresponding author were excluded from the analysis because many hospital staff are also working as faculty at schools and are supervising the research of students who may not work at the hospital. Further analysis was limited to research articles; other types of publications, such as reviews, editorial materials, letters, and corrections, were excluded from the analysis. The study also excluded case reports using medical record data from the analysis. For all research articles with TVGH staff as the primary author, a full-text PDF file of each article was downloaded.

For these research articles, the types of specialties, types of research, and impact factor were analyzed. Because the TVGH Department of Medicine and the Department of Surgery contain a wide range of subspecialties and a large number of staff, these departments were analyzed by subspecialties. Neurology and neurosurgery were also considered separately. The types of research were classified into basic research, database research, and clinical research.

To understand the conditions for the use of medical records, the collection was further analyzed for retrospective studies in clinical research. Each article was read in entirety to analyze the use and number of patients and the medical record items. The categories of medical record items were basic patient information (age, sex, height, weight, body mass index, and vital sign), notes (admission, progress, discharge, surgery, nursing, and procedure), medications, laboratory tests (blood, urine, stool, and body fluid analysis), and imaging reports (radiology, nuclear medicine, and pathology).

This study was conducted using descriptive statistics and Microsoft Excel (version 2016) software.

3. RESULTS

3.1. Characteristics of the Study Population

In 2018, a total of 966 articles indexed in Web of Science (SCI and SSCI) were published by the staff of TVGH, of which 808 (83.6%) were journal articles (Table 1). Of these 808 articles, 362 (37.5%) were first-authored by TVGH staff and 390 (40.4%) articles were by TVGH staff as the corresponding author. The Division of Cardiology ranked first in terms of the number of research articles written as the first author, with 36 articles, followed by Division of Gastroenterology and Hepatology (22), Department of Orthopedics (19), and Department of Radiology (19). The other departments with >10 publications in 2018 were the Department of Psychiatry, Neurological Institute–Neurosurgery, Department of Medical Research, Department of Obstetrics and Gynecology, Neurological Institute–Neurology, Department of Chest Medicine, Department of Pediatrics, and Department of Otolaryngology–Head and Neck Surgery (Fig. 1). The journals that published most

articles by TVGH staff were as follows: *Journal of the Chinese Medical Association*, *Scientific Reports*, *PLOS ONE*, *Medicine*, *International Journal of Environmental Research and Public Health*, and *Taiwanese Journal of Obstetrics and Gynecology*.

Among the 362 articles first-authored by TVGH staff, 219 (60.4%) were classified as clinical studies, 60 (16.6%) as basic studies, 53 (14.6%) as database studies, and 30 (8.2%) as other categories (Table 2). Other categories include internet data research, community research, medical device research, and medical education research. Retrospective research using TVGH medical records accounted for the largest proportion of clinical research articles (135 articles, 37.2%).

About 50% of the retrospective research using TVGH medical records had a case number <100 (67 cases, 49.6%) with an average number of 41 cases and 13 studies (9.6%) had a case number >1000 (Table 3). The article with the highest patient number is a study discussing use of emergency department resources in older adults, with 441,665 patients.[7] The article with the second highest patient number is one on prosthetic joint infection after total knee arthroplasty, using 10,768 cases.[8] The remaining articles were as follows: two articles used 5000–6000 cases, two used 4000–5000 cases, one used 3000–4000 cases, and six used 1000–2000 cases.

The analysis of the number of medical record items used in 135 retrospective research studies based on TVGH medical records showed that 118 (87.4%) studies used basic patient information, including gender and age (Table 4), whereas 4 (2.8%) studies used the basic patient information only. In addition to the basic information, notes written by professionals (e.g., surgical records and progress, procedures, and nursing notes) were used most frequently (73 cases, 54.0%), whereas medication information was used in 50 cases (37.0%), laboratory test data were used in 49 cases (36.2%), and height/weight/vital signs was used in 27 cases (20%).

4. DISCUSSION

Medical centers are responsible for teaching, research, and clinical care. Whether the center staff has published research articles has become one of the indicators of the quality of a hospital.[9,10] From this study, three important conclusions can be drawn. First, research based on medical records contributed to a substantial part of the publications by the hospital. Second, the majority of research based on medical records used of a small number of patients—no >100 patients. Third, in addition to basic patient information, documentation written by physicians and nurses in the wards or in the operation rooms played a major role in research based on medical records.

Previous studies have mainly focused on the type of journal in which the article was published, the number of citations, or the impact factor,[2] whereas few articles have discussed research patterns. Compared with the situation observed in the past, when observation and medical records alone were available, the pattern of today's research is quite different.[11] Seventeen case reports

Table 1

Articles published in 2018 with TVGH staff as first author or corresponding author

Document type	Total	First author	Corresponding author
Article	796 (82.4%)	357	384
Article; proceedings paper	11 (1.1%)	4	6
Article; book chapter	1 (0.1%)	1	
Review	65 (6.7%)	27	28
Letter	42 (4.3%)	25	31
Editorial material	40 (4.1%)	26	34
Correction	11 (1.1%)	4	3
Total	966	444 (46.0%)	486 (48.4%)
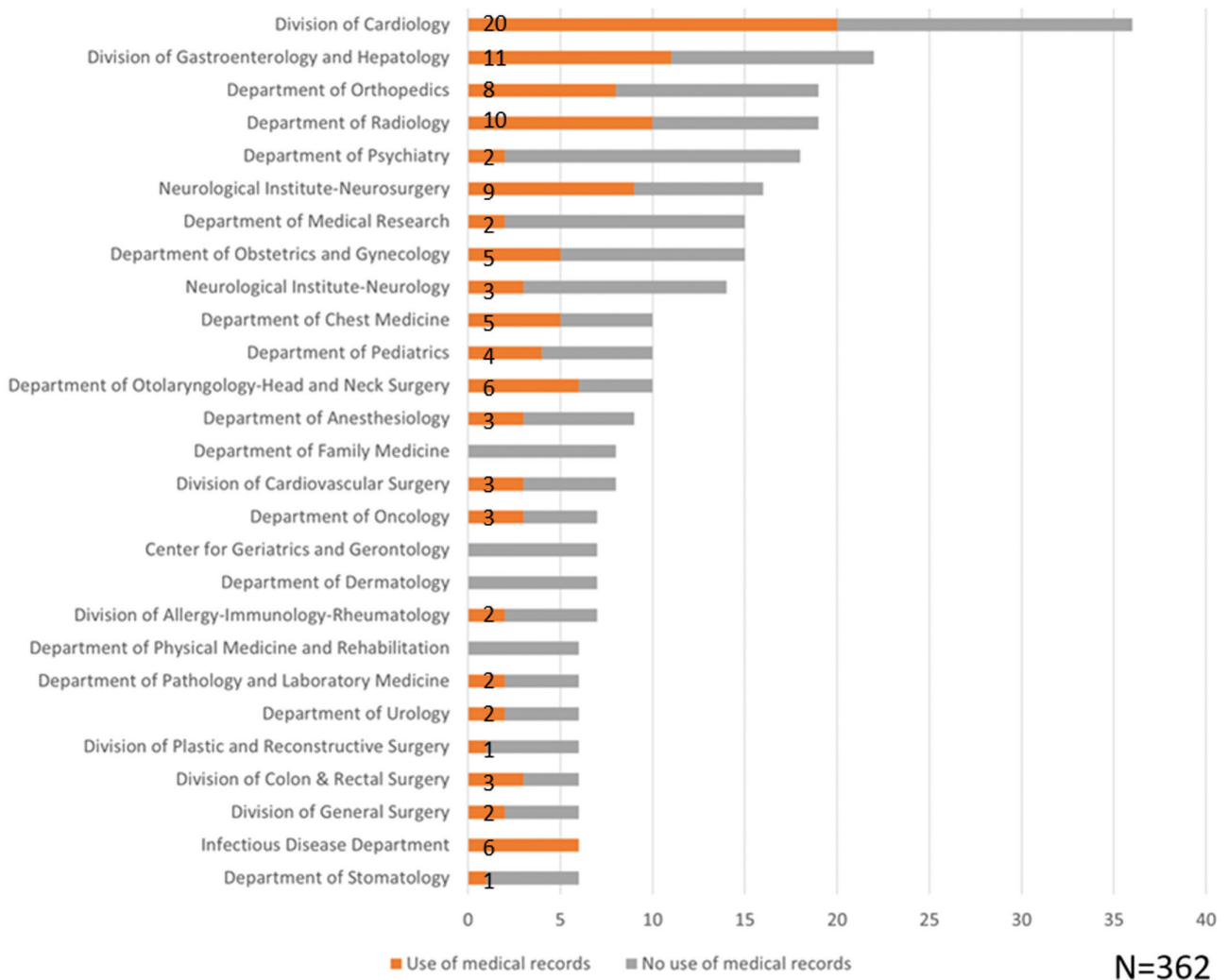
Fig. 1 Distribution of research published in 2018 by specialties at TVGH. This chart lists only those specialties with >5 articles. TVGH = Taipei Veterans General Hospital.

Table 2

Type distribution of articles published in 2018 by TVGH staff

Type of article	Number of articles	Impact factor (average)
Database research	53	4.231
Basic research	60	3.382
Clinical research	219	2.957
Prospective research	50	3.459
Retrospective research using hospital's medical records	135	2.791
Retrospective research using other hospitals' medical records	8	3.194
Retrospective research using cross-institutional medical records	5	2.905
Image analysis	3	8.380
Questionnaire research	1	2.293
Others	30	2.851
Total	362	3.120

TVGH = Taipei Veterans General Hospital.

Table 3

Number of cases in retrospective articles using hospital's medical records and prospective articles published in 2018 by TVGH staff

Case number	Retrospective articles using hospital's medical records	Impact factor (average)	Prospective articles	Impact factor (average)
≥1000	13	3.701	1	7.229
500–999	9	2.794	…	
300–499	11	3.473	3	4.140
100–299	35	2.667	17	3.307
≤100	67	2.566	24	3.297
Total	135	2.791	50	3.459

TVGH = Taipei Veterans General Hospital.

Table 4

Basic patient information in articles published in 2018 by TVGH staff

Items	Details	Number (%, n = 135)
Basic information	Gender, age	118 (87.4)
Notes	Admission, progress, operation, nursing, discharge	73 (54.0)
Medication		50 (37.0)
Laboratory	Blood, urine, stool, body fluid, etc.	49 (36.2)
Imaging report	Radiology, nuclear medicine, pathology	32 (23.7)
Body measurements	Hight, weight, vital signs	27 (20)

TVGH = Taipei Veterans General Hospital.

were published in 2018 by TVGH staff as first authors. The case report format has many limitations, such as the lack of epidemiological quantification, for the inability of performing causal inference, the absence of generalization, potential bias, overinterpretation, and emphasis on the rare versus common cases.[12] The number of case reports declined was reported in two references from the other countries.[13,14] Longer-term studies can be held in the future to assess the declination of case report.

Previously, medical records were handwritten or typed and collected in printed form.[15] Handwritten medical records have the disadvantage of being difficult to identify when conducting research, and paper medical records are not easy to organize and preserve.[16] With the advancement of technology and the widespread use of computers, medical records have been kept in digital form in Taiwan for many years.[17] Digital medical records have many benefits, such as being much easier for preservation, identification, exchange, and analysis.[18,19] The medical center for the current study, TVGH, was established >60 years ago and has accumulated a considerable amount of medical records. This huge database is a good foundation for research. In the era of electronic medical records, medical records are easier to obtain[20]; however, according to the results of the current study, the number of cases counted in most research articles is still <100. Obviously, most TVGH staff did not make efficient use of the available electronic medical records. Most of the TVGH staff still relies on the approach of manually reviewing each patient profile to collect data. Determining how to use the database to obtain data more efficiently and to increase the number of cases accessible is an important goal that necessitates further work.

In the current study, various notes were the most commonly used medical record items in addition to basic patient information. The types of medical records (excluding images) may be roughly classified as textual/numeric or structured/unstructured data. In this study, numeric laboratory data were used only in about 33% of research articles. Therefore, it can be seen that the records written by professionals are still very important for research. In the era of electronic medical records, it remains a daunting challenge to develop effective tools for text mining and natural language processing to extract and analyze data.[21] How to organize and analyze the descriptive and nonquantitative information is a future challenge.

This study had several limitations. First, letters as a type of article were not included in the study, although their content was often used in the medical record items; therefore, the results of the study may be underestimated. In the future, we will perform an analysis including all types of publications to assess the utilization of medical record items. Second, corresponding authors were excluded from the study because it was considered that they may be professors supervising students from another school or supervising residents from another hospital, and thus, the data used were not actually collected from TVGH. Thus, these results are also likely to be underestimated. Third, this study only counted the medical record items used in the SCI/SSCI accepted articles, and thus many domestic journals were not counted. This approach would underestimate the number and importance of medical record items.

In conclusion, this study summarized the current state of publications at an academic medical center in Taiwan in relation to the use of patient medical records. More than one-third of the publications utilized medical records; however, the number of patients and record items used was relatively limited. In the era of digitalization and large data analytics, the potential of medical records in research deserves attention. Investment in the establishment of a more accessible database of medical records to access nonstructural descriptive medical information should be considered.

REFERENCES

1. Collen MF. Clinical research databases–a historical review. *J Med Syst* 1990;**14**:323–44.
2. Madhugiri VS. Publication performance and research output of neurology and neurosurgery training institutes in India: a 5-year analysis. *Neurol India* 2015;**63**:338–46.
3. Coccia CT, Ausman JI. Is a case report an anecdote? In defense of personal observations in medicine. *Surg Neurol* 1987;**28**:111–3.
4. Soto CM, Kleinman KP, Simon SR. Quality and correlates of medical record documentation in the ambulatory care setting. *BMC Health Serv Res* 2002;**2**:22.
5. Ministry of Health and Welfare. Medical Care Act, 2020. Available at https://law.moj.gov.tw/ENG/LawClass/LawAll.aspx?pcode=L0020021. Accessed June 7, 2021.
6. Taipei Veterans General Hospital. Service statistics, 2019. Available at https://www.vghtpe.gov.tw/vghtpe/Fpage.action?muid=1487. Accessed June 7, 2021.
7. Chang JC, Yuan ZH, Lee IH, Hsu TF, How CK, Yen DH. Pattern of non-trauma emergency department resource utilization in older adults: an 8-year experience in Taiwan. *J Chin Med Assoc* 2018;**81**:552–8.
8. Wang FD, Wang YP, Chen CF, Chen HP. The incidence rate, trend and microbiological aetiology of prosthetic joint infection after total knee arthroplasty: a 13 years' experience from a tertiary medical center in Taiwan. *J Microbiol Immunol Infect* 2018;**51**:717–22.
9. Lopez-Escamez JA, Manrique-Huarte R, Martin-Sanz E, Trinidad G; Comisión de INVESTIGACIÓN de la SEORL-CCC. Analysis of scientific production in otolaryngology in Spain in the period 2011–2015. Acta Otorrinolaringol Esp 2018;**69**:275–82.
10. Blanc E, Boulieu R, Bertram D. A comparative evaluation of the publications of French public teaching hospitals from 2004 to 2014. *Bull Cancer* 2019;**106**:1094–103.
11. Beins BC. Research Method: A Tool for Life. 4th ed. Cambridge, UK: Cambridge University Press; 2018.
12. Zainal Z. Case study as a research method. Jurnal Kemanusiaan 2007;**5**:1–6.
13. Nissen T, Wynn R. The history of the case report: a selective review. *JRSM Open* 2014;**5**:2054270414523410.
14. Nissen T, Wynn R. The clinical case report: a review of its merits and limitations. *BMC Res Notes* 2014;**7**:264.
15. Reiser SJ. The clinical record in medicine. Part 2: reforming content and purpose. *Ann Intern Med* 1991;**114**:980–5.
16. Engle RL Jr. The evolution, uses, and present problems of the patient's medical record as exemplified by the records of the New York Hospital from 1793 to the present. *Trans Am Clin Climatol Assoc* 1991;**102**:182–9; 182–9.
17. Li YC, Yen JC, Chiu WT, Jian WS, Syed-Abdul S, Hsu MH. Building a national electronic medical record exchange system - experiences in Taiwan. *Comput Methods Programs Biomed* 2015;**121**:14–20.
18. Kim E, Rubinstein SM, Nead KT, Wojcieszynski AP, Gabriel PE, Warner JL. The evolving use of electronic health records (EHR) for research. *Semin Radiat Oncol* 2019;**29**:354–61.
19. Rakel RE, editor. Textbook of Family Medicine. 7th ed. Philadelphia: Saunders Elsevier; 2007.
20. Häyrinen K, Saranto K, Nykänen P. Definition, structure, content, use and impacts of electronic health records: a review of the research literature. *Int J Med Inform* 2008;**77**:291–304.
21. Chen X, Xie H, Wang FL, Liu Z, Xu J, Hao T. A bibliometric analysis of natural language processing in medical research. *BMC Med Inform Decis Mak* 2018;**18**(Suppl 1):14.