1. Cover Page

Final Performance Report

Grant Number: PW-228226-15

Title of Project: Medicine at Ground Level: State Medical Societies, State Medical Journals, and the

Development of American Medicine

Name of Project Director: Scott Podolsky

Grantee Institution: Harvard Medical School

Submission date: July 26, 2017

2. Narrative Description

a) Project Activities:

The rationale for "Medicine at Ground Level" was that state medical society journals document the transformation of American medicine at both the local and national level, serving as sites not only for scientific articles, but for medical talks, local news regarding the medical profession, pharmaceutical and device advertising, and unexpurgated musings on medicine and society throughout the 20th century.

Our initial plan was to digitize 117 titles from 48 states, from 1900 to 2000, comprising 2,500,369 pages in 3,579 volumes. In terms of content, the scope of the completed projected has exceeded even our initially ambitious plans. The five Medical Heritage Library ("MHL") partners in the project – the Center for the History of Medicine in the Francis A. Countway Library at Harvard Medical School, the New York Academy of Medicine, the College of Physicians of Philadelphia, the University of California, San Francisco, and the University of Maryland-Baltimore – to this point have digitized 97 titles (as certain titles were found to be shared across multiple states) from 48 states along with the District of Columbia and Puerto Rico, comprising 2,766,898 pages in 3,816 volumes.

Beyond the Internet Archive's (IA's) access portal through which MHL content is delivered, we developed a State Medical Society Journals portal in the IA at

http://www.medicalheritage.org/content/state-medical-society-journals/, connected to our main MHL presence, augmented by state-by-state links to the journals

at http://www.medicalheritage.org/content/state-medical-society-journals/journals-by-state/.

Moreover, the "Medicine at Ground Level" project spurred us (with the input of our MHL advisory board) to develop the MHL's own advanced search web portal

(http://mhl.countway.harvard.edu/search/) and search interface, making its 197,196 (as of July 2017)

items full-text searchable, complete with proximity search, date search, and faceted searching by contributor, languages, and collections (including the state medical journals as a "collection"), among other features.

To accomplish its objectives, the five partners met by phone nearly weekly throughout the project, meeting in person yearly at the annual meeting of the American Association of the History of Medicine (AAHM), and comparing metrics and workflows related to cataloging and digitization.

As proposed, the MHL is ensuring continued access and long-term preservation of its grant-funded digitized content by utilizing both the Internet Archive (IA) and individual institutional preservation repositories to sustain its digital files. Currently, IA has two geographically distributed locations for backing up copies of files (Egypt and Amsterdam), and is currently planning the Internet Archive Canada, which will serve as a complete backup of all IA content. This fall, the Center will be working with Harvard's Library Technology Services unit to programmatically acquire and ingest all of its MHL content to its Digital Repository Service (DRS); this scripting solution will be available to MHL members. While most grant partners are keeping copies of their digitized content in the IA and locally in digital preservation systems, the University of California, San Francisco will also be pushing its content to Hathi Trust. Finally, metadata for all of the MHL's IA content will be ingested into the Digital Public Library of America, where it will be visible to the public while also serving as a secondary metadata repository. The MHL is now working with the DPLA to ensure a "fresh harvest" of its content to the DPLA to reflect the availability of state medical society journals.

In terms of outreach, as described below, each participant notified its local audiences regarding the ongoing scope and utility of the project, while the whole group had the chance to meet with academic historians at the annual meeting of the AAHM in Minneapolis in 2016 and in Nashville in 2017. Finally, while the scope of the project exceeded the funding for digitization provided by the NEH grant, the

importance of the project and the nature of the team assembled for its completion led to our receiving additional funding from the Arcadia Fund through the Harvard University Library (\$60,000), and from Harvard Medical School through its John Talbott fund (\$53,000 to date). Moreover, Harvard Medical School contributed additional funding (\$20,000) over the course of the project for the completion of the MHL full-text search interface (enabling historians, clinicians, and researchers of all types to derive unique insights into the history of American medicine and public health.

b) Accomplishments

The primary accomplishment of the project has been the collaborative digitization of nearly every state medical journal in the United States, from their initiation almost to the present day. Given both the desire for partnership with the state medical societies (e.g., in promoting this rich resource to their constituent audiences) and the legal need to obtain permission to digitize in-copyright materials, this initially entailed multiple discussions with representatives (usually executive vice presidents or heads of publication, communication, or legal divisions, depending on the size of the society), prior to receipt of written approval for the digitization of in-copyright materials. We reached agreement with every medical society in the United States, with the exception of Massachusetts (which has already fully digitized and monetized the New England Journal of Medicine) and New Hampshire (which never had a medical journal apart from its onetime sponsorship of NEJM); and we were able to include the Medical Society of the District of Columbia and the Puerto Rico Medical Association (the latter reached out to us after one of its members heard of our efforts through our presentation at our 2016 American Association for the History of Medicine (AAHM) panel, "Medicine at the Ground Level from the Medical Heritage Library: State Medical Societies, State Medical Journals, and the Development of American Medicine and Society," chaired by PI Scott Podolsky and featuring as speakers historian Nancy Tomes (State University of New York, Stony Brook, archivist Polina Ilieva (with MHL project contributor

University of California, San Francisco), and current MHL co-chair Melissa Grafe (Yale University) For forty-five of the states, the District of Columbia, and Puerto Rico, we went back and obtained formal contracts enabling the perpetual noncommercial re-use of the materials (see Appendix 2 for a sample agreement). While our initial plan was only to digitize materials through 2000, three states were eager to have their post-2000 volumes added to the database; we accommodated this (and while not legally obligated to do so, we intend to continue to add volumes from such states moving forward). Ultimately, to date we have digitized 97 titles from 48 states along with the District of Columbia and Puerto Rico, comprising 2,766,898 pages in 3,816 volumes.

This exceeded the initial goals of the project, and was enabled through additional funding that the NEH grant itself enabled us to acquire. It is difficult to overstate the range of queries that can now be answered through the availability of such a remarkable database. It should be noted, though, that there were challenges in completing the digitization of the materials. By November 2016, as the majority of the volumes had been digitized, it became apparent that despite the vast holdings of the five participating libraries, there were still 236 volumes not held in acceptable condition to be digitized. Through working in WorldCat and in conversation with our librarian colleagues, we identified seven additional partners – Brown University, Columbia University, Emory University, the National Library of Medicine, the University of Minnesota, the University of North Carolina, and Stanford University – who could provide the "missing" volumes. The work of digitizing these last volumes has been ongoing and is expected to be completed no later than September 1, 2017, with the additional cost of digitization to be paid from Harvard Medical School's John Talbott fund.

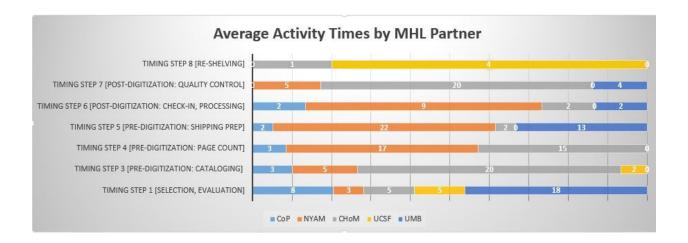
A second major accomplishment of the project has been the advancement of the search interfaces used to interact with not only the state medical journals, but with the MHL materials content as a whole. As hoped, even researchers without much training in the digital humanities search expertise can now

perform sophisticated queries to investigate health trends and outcomes over time and region. And for digital humanists, the full text and metadata is freely accessible for whatever search queries or visualizations they would like to design. One interesting development that has the potential to further extend the capabilities of the advanced tool is the work of Helge Holman and Vinay Goel, for whom the MHL provided researcher scenarios for making their ArchiveSpark web extraction tool extensible to the MHL's full text corpus (https://github.com/helgeho/MHLonArchiveSpark). This Git repository contains the required components for ArchiveSpark to work with Medical Heritage Library (MHL) collections; those comfortable with the recommended software environments can query the corpus with the search tool and then extract the full text for results and perform additional analysis, as illustrated in this example: https://github.com/helgeho/MHLonArchiveSpark/blob/master/examples/MhlPolioSymptomsSearch.ipynb

A third product of grant work includes the development of preliminary time and labor data for performing activities related to digitization in a collaborative environment. For the duration of the grant, partners agreed to track their work utilizing eight categories which covered the process from selection of the volume to re-shelving after digitization. These activities are described in the appended digitization workflow document and were captured by partners using the workflow document in spreadsheet form as a model for data collection. The below chart illustrates the timing per volume for the activities successfully tracked for the grant. (Note that Step 2, De-duplication, was specific to Countway's workflow only and was not included in the analysis.)

Overall, across partners, only the selection and evaluation of volumes for inclusion in the project could be captured by all partners (5/5), with steps 3, 5, and 6 captured by 4 out of 5 partners, steps 4 and 7 by 3 out of 5 partners, and step 8 captured by only 2 partners. The data on the one hand demonstrate the

variability that remains for conducting various activities, depending on local contextual factors. At the same time, they do provide general parameters that will be useful for future MHL digitization initiatives.



While we had initially hoped to provide useful conservation metrics on bringing fragile materials to "digitization readiness," it became readily apparent to us which materials could or could not be digitized without significant risk of further damage to the volumes. Moreover, we did not have the resources to conserve at-risk materials into materials capable of being safely digitized. We thus did not attempt to bring fragile materials to digitization readiness, but rather looked for alternative versions among our collections, or (as above) outside our MHL member collections for available alternatives.

Fourth, the project produced standardized metadata guidance (see Appendix 5) and clean-up. As the project developed, it became clear that metadata enrichment would be a necessary post-digitization activity not anticipated in the original process. Although metadata guidelines were developed as part of the grant, this guidance typically applied to the creation of new metadata for under-cataloged volumes; partners generally relied on existing catalog metadata for the project. Once metadata across collections became united in the Internet Archive, variations in practice across partners surfaced immediately and it became evident that project metadata required normalization in order to ensure maximum

discoverability across platforms, including the advanced search tool; the Center for the History of Medicine took on this task. As of June 2017, 6,513 changes have been made to state medical society volume metadata, including standardizing dates, creators, and subjects. With the help of the Internet Archive's Andrea Mills and Tim Bigelow, project assistant Darren Young and Center cataloger Joan Thomas were able to edit any record tagged *statemedicalsocietyjournals*. Tasks included: Adding subject headings, creators, and publishers; Adding dates to volume fields; Correcting the "v. volume" problem (IA supplies the word "volume" for each item, and CSV's that included "v." resulted in this redundancy; Correcting dates (IA's default setting is to take the date from the bibliographic record, which does not reflect the dates of individual volumes); Removing initial articles from titles in order to ensure that they sort correctly; Correcting spelling errors. One goal of the enrichment project was to standardize the display in the Internet Archive as much as possible; another was to ensure accuracy. During the active period of the grant, the Center spent over 500 hours correcting errors or enriching metadata, and the work continues.

Fifth, with respect to outreach, each of the contributing sites promoted the project to their local audiences. From the Center for the History of Medicine, see the following blog post:

https://cms.www.countway.harvard.edu/wp/?p=11178

From the New York Academy of Medicine, see the following blog post:

https://nyamcenterforhistory.org/2016/06/15/digitizing-medical-journals-of-state-societies/

From the College of Physicians, see the following blog posts:

- http://www.collegeofphysicians.org/histmed/text-mining-in-west-virginia/
- http://www.collegeofphysicians.org/histmed/tag/state-medical-society-journals/

From UCSF, see the following blog posts (they also spoke about the project at the Bay Area History of Medicine society, and UCSF Department of Anthropology, History and Social Medicine meetings):

- https://blogs.library.ucsf.edu/broughttolight/2015/09/11/joint-project-to-digitize-statemedical-society-journals-1900-2000-funded/
- https://blogs.library.ucsf.edu/broughttolight/2016/08/17/digitized-state-medical-journals-searching-alcohol-and-prohibition/
- https://blogs.library.ucsf.edu/broughttolight/2016/07/14/highlights-from-the-state-medical-journal-digitization-project/

The University of Maryland-Baltimore promoted the project in *The Elm*, its newsletter to its constituents.

In addition, the groups collaboratively promoted the materials to the Archivists and Librarians in the History of the Health Sciences (ALHHS) through its publication, *The Watermark*, in articles in both the Fall 2016 and Winter 2017 editions. The MHL's "Medicine at the Ground Level" session at the AAHM meeting in Minneapolis on April 30, 2016 was well-attended (over 40 audience members), and generated considerable enthusiasm about the utility of the project.

Importantly, with the project and search interface now nearly completed, we will soon be closing the loop and encouraging the participating state medical societies to share the links directly with their members (several of the state medical societies have already contacted us about this). This will be accomplished as soon as the projected is completed (i.e., approximately September 1, 2017).

c) Audiences

To this point, despite the brief time they have been made available and the provisional nature of the interfaces used to reach them, the state medical journals have been accessed 362,531 times. More broadly, the Medical Heritage Library's items have been downloaded 42,244,739 times. While our most obvious audience remains historians of medicine, public health, public policy, race, gender, and art, we expect that the materials will be salient to a wide array of scholars looking at the ongoing shaping and depiction of "healthy" individuals and the measures seemingly required to ensure such health. We also suspect that the approximately one million physicians in the United States will become an increasingly important user base for the collection, especially through the encouragement of their state medical societies.

With respect to the local impact of the state medical journal project on our local audiences and our own institutions, several members have described the increased access to the materials provided to their constituents, and in reinforcing the values of their libraries to their stakeholders. Colleagues in the history of medicine at Yale University have already begun incorporating the online state medical journal collection into their teaching.

During the grant period, the MHL's social media outlets have experienced significant growth: we now have 4,290 Twitter followers, a little over double since the start of this project. Over the course of the state medical journals project, our Project Coordinator put a number of images, mostly advertisements, from various journal volumes on our Twitter account for example,

see https://twitter.com/MedicalHeritage/status/867818216772513792/photo/1 from May 2017, which received 1,420 impressions (number of views on Twitter)

and https://twitter.com/MedicalHeritage/status/848987264797552640/photo/1 from April 2017, which received 926 impressions. We have also heavily promoted the project on the MHL's Facebook page (https://www.facebook.com/medicalheritagelibrary).

d) Evaluation

In addition to the workflow analysis performed on discrete digitization activities, several forms of evaluation were performed throughout the project. First, among ourselves, we collected and shared data concerning learned efficiencies, permitting ongoing feedback that helped to improve the conduct of the project.

Second, in our ongoing development of the MHL search interface, we and our MHL scholarly advisory committee serially tested the utility of the interface, using the state medical journal corpus as our test body. The results of such evaluation contributed to the surfacing of metadata issues and to by-hand volume-level metadata cleanup essential to the long-term utility of the project.

e) Continuation of the Project

"Medicine at Ground Level" has extended the reach of the MHL and has strengthened the collaborative partnership that is the MHL itself. We intend to continue to build on its success and expand the scope of the MHL itself. Several project members – including the Center for the History of Medicine, Countway Library, the College of Physicians, the New York Academy of Medicine, Yale University, the University of Minneapolis, and the University of California, San Franciscohave submitted a grant proposal to the Council on Library and Information Resources to catalog and scan approximately 11,540 previously unpublished public health pamphlets. The outcome of such a project – with items initially directed more towards the general public, and the public health professions – would beautifully complement the state medical journal project, whose content was ostensibly directed towards the medical profession per se.

We just found out (on 7/14/17) that we've advanced to the final round of the CLIR grant application process

Moreover, building on the success of "Medicine at the Ground Level," in partnership with MHL member

institutions, the University of California, San Francisco submitted an NSF RIDIR grant for the creation of analytical tools that would serve to enhance database use of historical patient data to address significant questions in the social behavioral and economic sciences. This grant proposal, if accepted, would enable the digitization of select nineteenth and twentieth century health records for the purpose of creating researcher access to content with protected health information without violating privacy.

Finally, "Medicine at the Ground Level" has demonstrated the utility – and the labor-intensive challenges – of digitizing in-copyright items. It requires investigating the legal owners of such items, and of demonstrating to them the utility of having their items digitized and made freely available. Several of the most widely read physician publications of the 20th century – *Modern Medicine*, *Medical Economics*, and *M.D.* – were "throwaway" journals that lived in physician waiting rooms and homes, shaping the ideals of medical practice. We would hope to digitize such widely read runs of journals, again making them full-text searchable (especially as their metadata has not been harvested by *Index Medicus* or its

f) Long Term Impact

successors like PubMed).

As described above, the grant has had multiple long-term effects on both our local institutions and with respect to the MHL itself. For example, at the Countway Library and Harvard, the NEH grant enabled us to procure \$60,000 from the Arcadia Fund (through the Harvard University Library), and \$53,000 through the John Talbott fund for the completion of the digitization project itself. It has enabled the Center for the History of Medicine to advocate for funding from the medical school and university for journal digitization more broadly.

At the University of Maryland-Baltimore, it motivated MedChi, The Maryland State Medical Society, to further digitize its own holdings, and the Health Sciences and Human Services Library to consider the digitization of dental and pharmacy society publications.

At the level of the MHL as a whole, "Medicine at Ground Level" positioned us to interact with potential new MHL partners (e.g., in the supply of volumes not held by the core grant libraries), and indeed directly led the recruitment of the University of Minnesota as a partner in the CLIR grant described above. And as above, it led to the further collaboration among the MHL partners that led to the formulation of its subsequent CLIR grant.

g) Grant Products

The primary product of the grant is the access to the state medical journals themselves. This is available through the Internet Archive site at:

http://www.medicalheritage.org/content/state-medical-society-journals/

A key complementary product of the grant is the full-text search portal and interface, available at: http://mhl.countway.harvard.edu/search/

3. APPENDICES

Appendix 1: Provisional list of journals included (please note that this will continue to be modified until the presently missing volumes have been digitized and remaining metadata issues have been reconciled):

Alabama

Alabama Medicine (1984-1996) Journal of the Medical Association of the State of Alabama (1931-1983)

Alaska

Alaska Medicine (1959-2000) Northwest Medicine (1903-1959)

Arizona

Western Journal of Medicine (1985-2000) Arizona Medicine (1944-1985) Southwestern Medicine (1917-1943)

Arkansas

Monthly Bulletin of the Arkansas Medical Society (1904-1906) Journal of the State Medical Society of Arkansas (1892-1895) Journal of the Arkansas Medical Society (1890-1892, 1907-2000)

California

Western Journal of Medicine (1974-2000) California Medicine (1946-1973) California and Western Medicine (1924-1946) California State Journal of Medicine (1902-1924)

Colorado

Rocky Mountain Medical Journal (1938-1979) Colorado Medicine (1903-1934, 1937, 1980-2000)

Connecticut

Connecticut Medicine (1958-2000) Connecticut State Medical Journal (1940-1957) Journal of the Connecticut State Medical Society (1936-1939)

Delaware

Delaware Medical Journal (1960-1995, 1997-2000) Atlantic Medical Journal (1923-1928) Delaware State Medical Journal (1912-1922, 1929-1933, 1935-1960)

District of Columbia

Medical Annals of the District of Columbia (1932-1974)

Bulletin of the Medical Society of the District of Columbia (1930-1931)

Washington Medical Annals (1902-1917)

Florida

Journal of the Florida Medical Association (1914-1971, 1973-1974, 1976-1978, 1981-1988)

Georgia

Journal of the Medical Association of Georgia (1912-1922, 1925-1992)

Hawaii

Hawaii Medical Journal and Inter-Island Nurses' Bulletin (1947-1961)

Hawaii Medical Journal (1941-1946, 1962-2001)

Idaho

Western Journal of Medicine (1975-2000)

Northwest Medicine (1903-1972)

Illinois

IMJ Illinois Medical Journal (1988)

IMJ, Illinois Medical Journal (1963-1987)

Illinois Medical Journal (1899-1962)

Indiana

Indiana Medicine (1984-1996)

Journal of the Indiana State Medical Association (1908-1980)

Iowa

Iowa Medicine (1984-2000)

Journal of the Iowa Medical Society (1961-1983)

Journal of the Iowa State Medical Society (1911-1928, 1933-1942, 1945-1959, 1961)

Kansas

Kansas Medicine (1985-1992, 1998)

Journal of the Kansas Medical Society (1903-1941, 1944-1984)

Kentucky

Journal of the Kentucky Medical Association (1965-1973, 1976-1979, 1981-1984, 1986-1999)

Journal of the Kentucky State Medical Association (1951-1954, 1956-1957, 1963)

Kentucky Medical Journal (1904-1950)

Bulletin of the Kentucky State Medical Association (1903-1904)

Louisiana

Journal of the Louisiana State Medical Society (1953-1963) New Orleans Medical and Surgical Journal (1844-1852, 1857-1859, 1866-1868, 1883-1952)

Maine

Maine Medical Journal (1930-1938) Journal of the Maine Medical Association (1911-1929, 1939-1970, 1972-1974, 1976-1980)

Maryland

Maryland State Medical Journal (1952-1964, 1974-1978, 1980-1984) Bulletin of the Medical and Chirurgical Faculty of Maryland (1908-1922) MMJ, Maryland Medical Journal (1877-1900)

Michigan

Michigan Medicine (1964-1968, 1970-1991) Journal of the Michigan State Medical Society (1902-1963)

Minnesota

Minnesota Medicine (1918-2000)

Journal-lancet (1912-1919, 1922-1923, 1925, 1928-1964, 1966-1968)

Journal of the Minnesota State Medical Society and Northwestern lancet (1905-1911)

Northwestern Lancet (1897-1905)

Mississippi

Journal of the Mississippi State Medical Association (1960-1961, 1963-1997) Mississippi Doctor (1928-1959) Mississippi Medical Monthly (1906-1914)

Missouri

Missouri Medicine (1953-1985, 1987-2000) Journal of the Missouri State Medical Association (1904-1952)

Nebraska

Nebraska Medical Journal (1971-1996) Nebraska State Medical Journal (1916-1952, 1954-1971)

Nevada

Western Journal of Medicine (1974-2000) California and Western Medicine (1924-1946)

New Jersey

New Jersey Medicine (1986-1987, 1989-1990, 1993-1995, 1997, 1999-2000) Journal of the Medical Society of New Jersey (1904-1914, 1916-1985)

New Mexico

Western Journal of Medicine (1980-2000)

Southwestern Medicine (1917-1972)

New Mexico Medical Journal (1911-1916)

The Journal of the New Mexico Medical Society (1905-1910)

New York

New York State Journal of Medicine (1901-1993)

North Carolina

Carolina Medical Journal (1900-1908)

North Carolina Medical Journal (1878-1899, 1978-1991)

North Dakota

Journal-lancet (1912-1919, 1922-1923, 1925, 1928-1964, 1966-1968)

Ohio

Ohio Medicine (1987-2000)

Ohio State Medical Journal (1905-1957, 1959-1986)

Oklahoma

Journal of the Oklahoma State Medical Association (1909-1914, 1917-1924, 1926, 1934-2000)

Pennsylvania

Pennsylvania Medicine (1966-1970, 1972-1973, 1975-1976, 1979, 1981-1991, 1993, 1995-1997)

Pennsylvania Medical Journal (1929-1965)

Atlantic Medical Journal (1923-1928)

Pennsylvania Medical Journal (1897-1922)

Puerto Rico

Boletín de la Asociación Médica de Puerto Rico (1903-1907, 1931-1954, 1959, 1961, 1963-1991)

Rhode Island

Medicine and Health (1996-2000)

Rhode Island Medicine (1992-1995)

Rhode Island Medical Journal (1917-1918, 1920-1982, 1984-1991)

South Carolina

Journal of the South Carolina Medical Association (1905-1920, 1922-1949, 1951-1961, 1963, 1965-2000)

South Dakota

South Dakota Journal of Medicine (1948-2000) Journal-lancet (1912-1919, 1922-1923, 1925, 1928-1964)

Tennessee

Tennessee Medicine (1996-1999)

Journal of the Tennessee Medical Association (1963-1996)

Journal of the Tennessee State Medical Association (1908-1909, 1911-1963)

Texas

Texas Medicine (1966-2000)

Texas State Journal of Medicine (1905-1966)

Utah

Western Journal of Medicine (1979-2000)

Utah State Medical Bulletin (1936-1937)

California and Western Medicine (1927-1932)

Vermont

Vermont Medicine (1916-1918)

Vermont Medical Monthly (1895-1914)

Virginia

Virginia Medical Quarterly: VMQ (1990-1992)

Virginia Medical (1977-1990)

Virginia Medical Monthly (1918-1962, 1965, 1972-1973, 1976)

Virginia Medical Semi-monthly (1896-1918)

Virginia Medical Monthly (1874-1896)

West Virginia

West Virginia Medical Journal (1906-1927, 1929-2000)

Wisconsin

Wisconsin Medical Journal (1903-2000)

Wyoming

Western Journal of Medicine (1983-2000)

Rocky Mountain Medical Journal (1938-1979)

Colorado Medicine (1926-1937)

Appendix 2: Sample Agreement (taken from agreement with the Medical Association of Georgia):

AGREEMENT

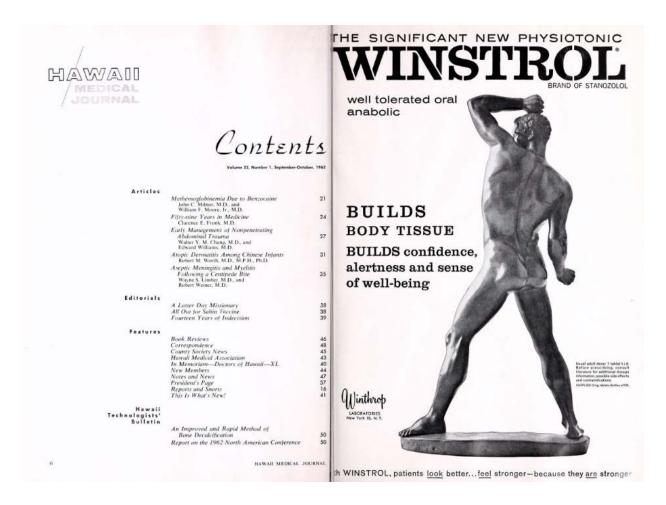
This Agreement (the "Agreement") is constituted between the Medical Association of Georgia ("the Society") and the Medical Heritage Library ("MHL"). The MHL is a group of collaborating libraries dedicated to enabling access to historical medical resources. The agreement applies to present and future iterations of the Society and the MHL.

The Society grants to the MHL a non-exclusive, worldwide license to the *Journal of the Medical Association of Georgia* (the "Licensed Materials"), including the right to digitize and make copies freely available online in perpetuity, for non-commercial uses, under the terms of a Creative Commons Attribution-Non Commercial 4.0 international License (CC BY-NC 4.0) or other emergent license with a different name, but governed by the same terms. The terms of the license require that the Society be attributed as the publisher of the Licensed Materials.

The Society states that it has the authority to license the Materials described in this Agreement.

The MHL acknowledges that the copyright and title to the marks relating thereto remain with the Society.	Licensed Materials and any trademarks or service
Scott H. Podolsky, M.D.	Date signed
Director, Center for the History of Medicine,	
Francis A. Countway Library of Medicine	
Chair, Scholarly Advisory Committee,	
Medical Heritage Library	
Authorized Signature for the Society	Date signed

Appendix 3: Sample pages:



From: Hawaii Medical Journal 22 (September-October 1962)

The result of the acquisition of new knowledge—as an aftermath of the war—progress and interest in the field of mental diseases has been anazingly beneficial to the medical practitioner, as well as those he served. The revent advances in pychicity and the study of the patient is a consulption of the progress and interest in the field of practitioner, as well as those he served, of the patient is a consulption of the progress and progress study of the progress study of the progress of the p

January, 1948

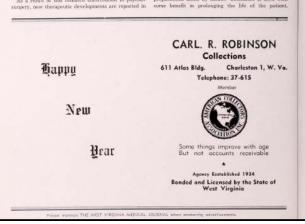
THE WEST VIRGINIA MEDICAL JOURNAL

X-rays, however, cannot always be used, particularly in persons with acute leukemia.

The disease is practically always field, and treatment at present does little mere than stave off the evaluation of \$1,000. Only a small part of this rise has come about through the increase in the child population, intended as a bettermined of the continues such as bettermined to which are not known, that research most continue and research strough animal experimentation in the National Society for Medicial Research, members of which are now earrying on research in various medical schools and laboratories throughout the country—Health Talk, Education Comm. Illinois St. Med. Sec.

INCREASE IN POLIO CRIPPLIS

Poliomyelish has attacked an unusuably large number of persons in the United States in recent years. In face, the four years from 19th drough 19th constitute the longest period of sustained high incidence of the disconse in the history of the country. In 19th slower, the the longest period of sustained high incidence of the disconse in the history of the country. In 19th slower, there were more than 25,000 cases, a figure which was exceeded only once before—in 1019, when more than 27, and 19th of the purpose of the country in the first of the country late of January 1st of the current year, it is further estimated that there were nough 74,000 persons under a go 21 who were crippled to some degree by the disease; at McKeal Bulletin.





WHEELING CLINIC WHEELING, WEST VIRGINIA

Internal Medicine:

Clinical Leharceton, B.S.

ADMINISTRATION

From: West Virginia Medical Journal 44 (January 1948)

Step 1

Tracking/Timing Category: Evaluation

Scope: Title by title examination and evaluation of journals

Notes: Step 1 generates data that goes into the journals spreadsheet [*]

- Examine volume: is it digitizable?
- If rejected, return to shelf. Enter reason(s) for rejection in spreadsheet Columns I-M and Column O
- Does the assigned volume contain advertisements? Enter Yes/No in spreadsheet
 Column H*
- Does the volume contain fold out advertisements? Enter Yes/No in spreadsheet Column
 [?] and record where in the Notes field [Column __0_]
- Does the volume contain advertisements that "bleed" out in to volume gutter? Enter Yes/No in Column [P]
- Enter final decision re: whether it can be scanned in spreadsheet Column N* (Y/N)

Step 2 (For Countway only)

Tracking/Timing Category: De-duplication

Scope: Activities to generate list of titles to be assigned to collaborators for scanning

- Combine collaborator spreadsheets to create master evaluation document
- Identify gaps in title holdings
- Review volume condition notes, flagging volume in the best condition for digitization
- Create a list of scanning assignments for each collaborator and review findings as a group
- Update scanning assignments as necessary

Step 3

Tracking/Timing Category: Cataloging

Scope: Verify and/or update catalog records for IA metadata compliance **Notes:** Step 3 generates data that goes into the journals spreadsheet [*]

- Do any cataloging necessary to meet local or MHL minimum metadata standards
- Track timing by title*

Step 4

Tracking/Timing Category: Page Counting

Scope: Generate page counts for volumes to be scanned

Notes: Step 4 generates data that goes into the journals spreadsheet [*]

- Make a tally of pages per assigned volume
- Record page count per assigned volume on the spreadsheet Column E*

Step 5

<u>Tracking/Timing Category:</u> Digitization, Pre: Shipping Preparations

Scope: Create local tracking paperwork and prepare volumes for shipment

Notes: Step 5 generates data that goes into the journals spreadsheet

- · Generate any local reports necessary for tracking
- Generate Partner MetaApp report
- Email Partner MetaApp report to IA
- Generate any necessary local moving requests or documentation
- Prepare carts or boxes of volumes for transfer to scanning site

The next two steps can be done in either order (6, 7 or 7, 6)

Step 6

Tracking/Timing Category: Digitization, Post: Check-in and Processing

Scope: Unpack digitized volumes and inspect for physical damage

Notes: Step 6 generates data that goes into the journals spreadsheet

- Remove returned volumes from shipping materials
- Check returned volumes against Partner MetaApp report
- Check for physical damage to returned volumes
 - If damage is found, report to IA scanning center

Step 7

Tracking/Timing Category: Digitization, Post: Quality Control

Scope: Audit digitized versions of volumes, verify metadata, and compile list of volumes requiring rescans

Notes: Step 7 generates data that goes into the journals spreadsheet

- Start by pulling every 5th volume to check quality; if quality of images proves good, revise the number upwards: every 15th volume, etc.
- Check each page of the volume against the scanned image
- Check metadata for the volume
- Make sure full text OCR files are present
- Make sure all file formats are viewable
- If any of these conditions are **not** met, set the volume aside
- For problem volumes:
 - Send link of problem page to IA
 - Agree upon solution
 - Print out email re: resolution for local filing
 - Return any volumes that require re-scanning to the appropriate work area

Step 8

Tracking/Timing Category: Re-Shelving

Scope: Return scanned volumes to stacks or storage

Notes: Step 8 generates data that goes into the journals spreadsheet [*]

• Return volumes to their original shelves*

Appendix 5: Metadata Guidance:

Using the Internet Archive's Partner Meta App for Serials

When digitizing items with the Internet Archive (IA), institutions can use the Partner Meta App to create the metadata that accompanies each digitized item on archive.org. To use the Partner Meta App, institutions only need to fill out a simple spreadsheet, the instructions for which are available at https://archive.org/details/PartnerMetaApp. However, when digitizing serials specifically, there are a few additional things to keep in mind. Along with the columns IA specifies in its instructions, it is necessary to also include a DATE column. Along with the VOLUME column, the DATE column is necessary to accurately and consistently format the enumeration and chronology designations of each serial volume, and for those volumes to display in an orderly and understandable way.

Typically, the Partner Meta App uses a Z39.5 connection to the contributing library's catalog to pull in the bibliographic data that displays with each item on archive.org. However, a conflict arises due to the fact that IA creates a separate record for each physical volume it digitizes, while library catalogs typically use a single record for all volumes of a serial. Therefore, for each physical volume, "enum/chron" data must be supplied in the VOLUME column of the spreadsheet because it cannot be pulled from the original bibliographic record via the Z39.5 connection.

In addition, the Z39.5 connection will pull date data from the bibliographic record that will reflect the serial title as a whole, not the individual volume being digitized. This results in messy metadata in the IA record, so it is necessary to supply date data appropriate to the individual volume in the DATE column and override the data that it pulls from the bibliographic record.²

Below is an example showing a spreadsheet filled out for the Maryland Medical Journal:

¹ Please use the accompanying spreadsheet template.

² The IA date field will only allow single dates, not date ranges, so only single dates are allowed in the DATE column. Date ranges are, however, allowed in the volume field.

	Α	В	С	D	Е	F	G	Н
1	search_id	title	volume	date	language			
2	136154	Maryland medical journal	v.39, (1898)	1898	eng			
3	136154	Maryland medical journal	v.54, (1911)	1911	eng			
4	136154	Maryland medical journal	v.61, (1918)	1918	eng			
5								
-								

The above would create a record in IA for each volume that is digitized. Below is an example record for Maryland Medical Journal, v.54, (1911). The volume information and the publication date (next to the title) were supplied in the spreadsheet while the rest of the data, such as Subject and Publisher, was pulled from the bibliographic record via the Z39.5 link.

Maryland medical journal (1911)

Volume: v.54, (1911)

Subject: Medicine; Medicine Publisher: [Baltimore, Md. : s.n.]

Language: English Call number: 136154

Digitizing sponsor: Open Knowledge Commons and the National Endowment for the

Humanities

Book contributor: The College of Physicians of Philadelphia Historical Medical Library

Collection: medicalheritagelibrary; collegeofphysiciansofphiladelphia; americana

Full catalog record: MARCXML

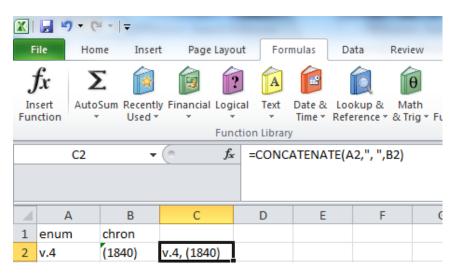
Description

Subtitle varies: A semi-monthly journal of medicine and surgery, 1881-1882; A weekly journal of medicine and surgery, 1883-1899; A journal of medicine and surgery, 1900-1918

Includes Proceedings of the Medical and Chirurgical Faculty of the State of Maryland, 1905-1908

Manually filling out the spreadsheet field-by-field can be a time-consuming process, but there are a few techniques to speed it along. If you have the ability to pull fields from selected MARC records into an Excel spreadsheet, you can then copy and paste the volume and date data along with title, etc. into the spreadsheet.

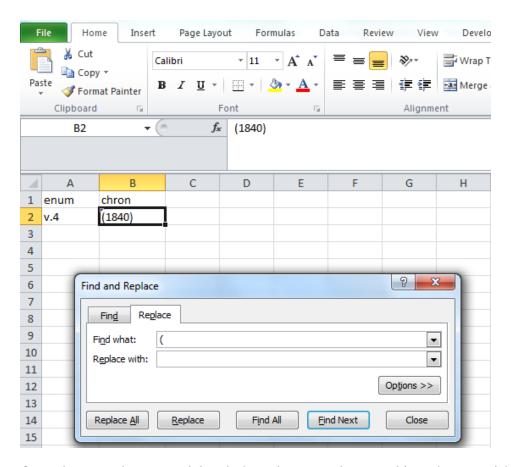
Before pasting data into the VOLUME and DATE column of the spreadsheet, first make sure that it is accurate as well as consistently formatted.³ After that, it may require a combination of using the Excel function, "Concatenate", and "Find and Replace" to further format the data for pasting into the spreadsheet. For instance, if you pull the "enum/chron" data into a spreadsheet and it is separated into two columns, you can use the "Concatenate" function to create a single column to copy and paste⁴ into the VOLUME column of the spreadsheet:



We could then just copy the Chron column and paste it into the DATE column of the spreadsheet, but we do not want the dates formatted with parentheses in the DATE column. To remove the parentheses, we can use "Find and Replace" to replace each parenthesis with a blank and then copy and paste the result into the spreadsheet. It should be mentioned that when reformatting the data in different ways like this, it can be helpful to create a new excel workbook for each reformat:

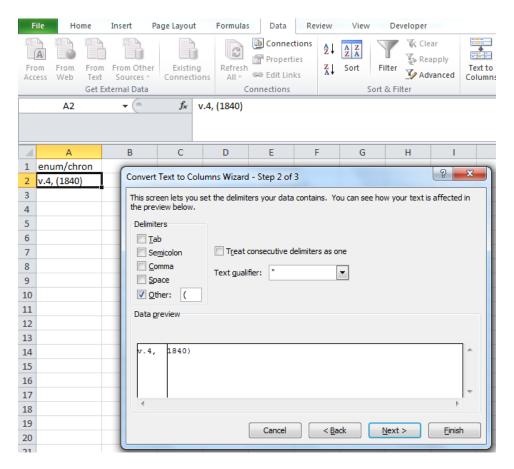
³ For guidelines for transcribing enum/chron data, see appendix.

⁴ When pasting the results of a function into another column, it is necessary to paste the value of the function and not the function itself.

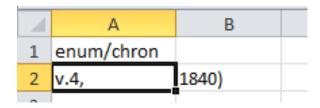


Once the parentheses are deleted, the column can be pasted into the spreadsheet DATE column. Be sure that your DATE data lines up with your VOLUME data.

If you pull the "enum/chron" data from a MARC record and it is *not* separated into distinct columns like in the example above, you can usually paste this data into the VOLUME column with minimal reformatting. However, you will still need to isolate the date data for pasting into the DATE column. For this, you can use the Excel feature, "Text to Columns". Starting with "enum/chron" in a format such as, v.4, (1840-1841), you can choose the parenthesis character as the delimiter to isolate the date:



It separates the original data into two separate fields:



After removing the final parenthesis, the resulting data can be pasted into DATE column of the spreadsheet.

In addition, if you do not have the ability to pull fields from MARC records into an Excel spreadsheet, you can also use Excel's "Auto Fill" feature to create a list of consecutive volume numbers and years. The results can be concatenated and reformatted for pasting into the spreadsheet.

These are just a few examples of how to automate the task of filling the VOLUME and DATE columns in the spreadsheet. The method may vary depending on the particular title and local catalog. Whatever method is used, if the VOLUME and DATE are formatted in the spreadsheet consistently and accurately, the IA record for each digitized volume can be displayed in an orderly and understandable way.

Sub-appendix: Suggested guidelines for transcribing enum/chron information into the volume column (from Columbia University)

- No spaces within enumeration.
- No spaces within chronology.
- Enumeration followed by comma.
- Chronology always in parentheses.
- v.1, (1991)
- (1996) /[title published without enumeration]/
- v.1, (2003-2004) /[all issues for one volume bound together, but chronology does not correspond to a single calendar year]/
- /v.3:no.1-4, (1966:Jan.-Apr.)
- v.3:no.5-8, (1966:May-Aug.) [/details of lower-level enumeration and chronology included only when higher level is bound in more than one piece]
- v.1/2, (1991/1992) /[published as a single combined issue by the publisher]/
- v.1-2, (1991-1992) /[two pieces separately published and bound together by the library]/
- no.1-4,no.6, (2001-2004,2006) /[bound volume lacks no. 5]/
- v.1/2, (1999/2000) Index [/index bound separately]/