Lister Hill National Center for Biomedical Communications

Clement J. McDonald, MD
Director

The Lister Hill National Center for Biomedical Communications (LHC) is an intramural research and development division of the NLM. To fulfill its mission, LHC:
- captures, processes, distributes, and uses high quality medical imaging data;
- develops and promotes health information technology standards to facilitate their adoption and meaningful use in electronic health records (EHRs), public health, and research;
- conducts research and development in biomedical natural language processing;
- develops mobile health and emergency response tools;
- provides health-related information to enhance patient engagement; and
- trains the next generation of medical informaticians.

Biomedical imaging research
LHC continued advanced imaging research to expedite diagnosis and treatment of diseases, support disease research, and advance 3D image modeling. (See "Patient Care Goes Mobile: Diagnosis through Image Analysis.")

A National Cancer Institute (NCI) cervical cancer biopsy study used LHC's Boundary Marking Tool to collect and annotate multiple-biopsy cervix images. (Single biopsy during colposcopy may miss 30-50% of high-grade cervix disease.) The study determined that multiple biopsies improve sensitivity of high-grade disease detection.

LHC used machine learning to improve its cervical cancer imaging tools. Staff used a support vector machine model to classify the disease severity in uterine cervix digital histology images. Performance accuracy was 95% for determining disease grade (Cervical Intraepithelial Neoplasia (CIN) 1-3 or Normal) and 88% for determining exact class (Normal, CIN1, CIN2, or CIN3). Automating classification can lessen the diagnostic burden on expert pathologists; provide a second opinion, potentially reducing inter-observer variability; and classify images more quickly.

The Open-*i*[SM] database of biomedical images enables search and retrieval of abstracts and images (including charts, graphs, and clinical images) from the open source literature and selected biomedical image collections. In FY2015, LHC added a special collection: the Orthopedic Surgical Anatomy Teaching Collection from the University of Southern California Digital Library.

LHC continued to support and update the Insight Segmentation and Registration Toolkit (ITK), an open source image analysis software that provides segmentation and registration algorithms in two, three, and more dimensions.

Health IT terminology and standards
In FY2015, LHC collaborations to facilitate health information exchange and meaningful use EHRs included:
- National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) to develop and publish guidance for using EHRs to improve chronic disease care;
- Centers for Medicare & Medicaid Services (CMS) to standardize their data elements across forms;
- nationwide efforts to standardize the reporting of newborn screening; and
- many others to develop and encourage the adoption of Common Data Elements. (See "Standardizing Big Health Data for Patient Care and Clinical Research.")

LHC organized agreements between LOINC® and IEEE, and between LOINC and RadLex (RSNA) to unify their respective coding systems and distribute mappings. LHC also worked closely with the FDA, CDC, and ONC to encourage the use of clinical standards—developed and supported by NLM—in instrument labels and drug trial submissions.

LHC enhanced LOINC's coverage, utility, and globalization and added new lab tests, such as maternal cell-free DNA tests for fetal trisomy, and genetic stool tests to predict colon cancer risk.

LHC developed SNOMED CT mapping resources to support the use of SNOMED CT® for coding diagnoses with greater clinical specificity than enabled by the International Classification of Diseases (ICD) codes (which were primarily developed and used for statistical and administrative purposes). LHC published a new release of the SNOMED CT to ICD-10 map, and Partners Healthcare System deployed the map in its clinical systems.

To help implement SNOMED CT, LHC extended its SNOMED CT CORE Problem List Subset, which covers 80% of commonly used terms and 84% of usage in a new problem list vocabulary. Staff also analyzed coverage of rare disease names in standard terminologies.

The RxNav browser and its APIs—for navigating RxNorm and other NLM-developed medication information resources—received one billion queries in 2015. Staff worked with OHDSI and PCORnet to map NDCs (historical drug codes) to RxNorm, extended the coverage of phenotypes in SNOMED CT through post-coordination, and added a new feature to RxClass to compute the similarity between two classes.

To enable standards-based EHRs and mobile health apps, LHC developed NLM-Forms, a suite of software tools to help build input forms and to readily populate form fields with terms from open medical vocabularies. LHC published the code and documentation for this open source software on GitHub.

Natural language processing
LHC staff worked with the Library Operations Division to refine and enhance two ongoing projects using natural language processing:
- The Consumer Health Information Question Answering (CHIQA) system helps staff handle the large volume of incoming customer service questions by creating sets of question types, recognizing spatial relations in questions, translating expressions to UMLS terms, correcting misspellings, and finding relevant information using the focus and type of question.
- The Medical Text Indexer (MTI) generates recommended MeSH terms for journal articles, allowing indexing work to proceed more quickly and efficiently.

Staff developed MeSH on Demand, an easier way to find Medical Subject Headings (MeSH) based on text processing, to help teachers annotate curriculum and class notes and to help authors find keywords for papers. LHC also conducted research to help Native Americans and other people with health disparities better utilize Semantic MEDLINE, a free tool that uses natural language processing and graph analysis to guide users to relevant MEDLINE citations.

LHC's Indexing Initiative worked with the FDA to determine the best MetaMap settings to extract information about side effects from Structured Product Labels. Through the PubMed Early Alerts pilot project, staff provided FDA safety officers with timely information about specific adverse drug events. Based on that success, LHC expanded the service to eight other FDA teams. Staff also deployed the NLM Scrubber de-identification software for the NCI SEER (Surveillance, Epidemiology, and End Results) Program. This effort improved the value of these data for research while preserving patient privacy.

Mobile health and emergency response
In collaboration with the National Institute of Neurological Disorders and Stroke (NINDS) Human Motor Control Section, NLM developed a mobile app to help people with Parkinson's disease record their daily health status. The app helps patients track their symptoms, including dyskinesia, tremors, falls, speech, mood, and medications. Such information can provide doctors with a fuller picture of their patients' condition to tailor treatments, adjust medications, and improve overall patient care.

LHC's Lost Person Finder Project helps reunite families after disasters. PEOPLE LOCATOR® was activated in response to emergencies including earthquakes, typhoons, and a volcanic eruption. Staff demonstrated TriageTrak, a hospital use-case version of the PEOPLE LOCATOR software, at the National Association of City and County Health Organizations 2015 Preparedness Summit.

Digital preservation
In collaboration with NLM's History of Medicine Division, LHC enhanced public resources for digital

preservation and history of medicine via *Profiles in Science* and *Turning the Pages*.

LHC published five new illustrated interviews in the *Conversations with Medical Informatics Pioneers* oral history collection: Donald A.B. Lindberg, G. Octo Barnett, Nina Matheson, Homer R. Warner, Sr., and Don Detmer.

Training

Our Medical Informatics Training Program hosted nine postdoctoral fellows in 2015 (seven PhD, one MD, and one MD/PhD), as well as students from NLM-sponsored university training programs, medical school, college, and high school. Students represented 17 US states and 11 countries, including the Philippines, Brazil, Nepal, and Hungary. Each participant spends between a few months and several years working on a research project under a mentor's guidance.