AMERICAN MEDICAL WRITERS ASSOCIATION AMWA JOURNAL V37 N4 / 2022

THEME ARTICLE

Technology to Further Medical Writing: Status and Future Vision

Helle Gawrylewski, MA[1] **and Nimita Limaye, PhD**[2]**/** [1]Former Senior Director, Global Regulatory Writing, Johnson & Johnson, New Hope, PA; [2]Vice President, Research, IDC Health Insights, Needham, MA

The panel members respond to questions that have been raised during sessions on different developments in using technology in medical writing. Medical writing is a profession dedicated to transforming data and analyses into useful and digestible information, whether that information involves regulatory applications or documents for the public and, specifically, patients. Decisions about treatments or granting of approvals depend on the distillation being accurate, clear, and understandable. Using available technology well can support this goal in that it is a means for shifting a writer's focus from what can be accomplished by artificial intelligence (AI), machine learning, and functions to the creation of content.

What are the drivers for using aspects of automation in medical writing, what gaps has it filled?

Helle Gawrylewski: In the pharmaceutical industry and in health care, automation in the writing process has spotty adoption depending on the size and digital sophistication of a company. Automations have been used effectively in writing by templates in which sections are prepopulated based on text from other documents. The protocol might be populated with text from the investigator's brochure (IB) or protocol concept document. Some companies have developed or acquired systems that can be used to accomplish this type of text before population and reuse. Microsoft Word itself has some slick automated capabilities that may not be fully used, like text tagging for reuse in other parts of a document. Another gap filled is the writing of routine text in documents like the safety narratives in clinical study reports (CSRs). Narratives are required in the CSR but are onerous to write, especially in cases in which there are many variables or many study participants with adverse events as in an oncology study. US Food and Drug Administration (FDA) reviewers have not been fond of safety narratives being totally written by automation, so this is not as common as it might be. But hybrid narratives, in which the data appear in brief tables and the discussion and assessment are written by a medical writer, can be efficient and accurate and medically useful. Safety narratives written entirely by AI require a large data set to teach the algorithms to produce adequate text.

The writing process also has benefitted from automation in review tools and quality control (QC). It's useful for the applicable style manual to be digitally available and automatically applied for document checking. Routine checks can be more efficient this way and a time-saver for the writer. Tools for the review cycle have also been used because it's tedious to send out sequential versions for document review when this can be done by a tool like *Please Review* and others, in which all comments can be seen by the team, tracked, and ultimately incorporated. Technology improves the process immensely and has had a positive impact not only on efficiency but also quality.

Other parts of an eCTD (electronic common technical document) have also benefitted by making the integrated summaries of safety and effectiveness (ISS and ISE) linked to the individual reports for a population, and the literature summaries can be captured by AI technology. I'm not sure how many companies take advantage of AI in this respect, but Nimita can perhaps address this more fully.

Other options for the use of AI and deep learning can be technical summaries of results for registries, and these can be populated when a CSR is written, as can the FDA Study Snapshots for safety by demographic characteristics that are required at approval. It's possible to populate the requirements automatically as an application is being built.

Scientific writing in another language, also referred to as translation or localization, benefits from at least some aspects of machine translation. Companies that use translation memories, machine learning, or advanced deep learning methods (also known as deep structured learning, with multiple layers between the input and output layers) can produce complex documents in many languages quickly, required for Lay Summaries in the European Union (EU) portal (implemented in January 2022). This

ISSN 2163-5315

type of automation requires standardization of concepts and terms so that coding can be used for digital exchange. Groups like the Clinical Data Interchange Standards Consortium (CDISC) and the Medical Dictionary for Regulatory Activities (MedDRA) code research terms and adverse events so they can be easily exchanged.

All of these uses require standardization of terms and definitions. A concept that assists in reusing information: text must also be considered data. Written content is data, and a document is just a compilation of data elements. Computer systems can be designed to use natural language processing (NLP) to understand written text. Machine learning and deep learning keep advancing, making these tools a substantial efficiency gain for any organization. It's also a boon to medical writers who can use the tools to summarize large amounts of data to ensure that all applicable resources are considered.

Nimita Limaye: Helle has made some great points. The future of medical writing is really about automation with the human in the loop. It is about leveraging not only robotic process automation and AI, but also about the use of machine learning (ML) techniques, such as NLP (which turns text into structured data) and natural language generation (which turns structured data into text). The challenge with training ML algorithms is the availability of massive labeled data sets. Transformer-based neural network architectures operate in a two-stage process, unsupervised learning on large volumes of unlabeled datasets, and then supervised learning on smaller amounts of labeled data. These are very powerful models and can be game changers, but these are still early days. There has been a very interesting report in the June 2022 edition of *Scientific American* about how a GPT-3 transformer was trained to write an academic paper about itself.

Automation will bring in significant efficiencies and reduce not only costs, but will also reduce the monotony associated with authoring the often-repetitive sections associated with regulatory documents and will improve quality. One is seeing a flurry of innovation, with technology vendors actively innovating to drive "intelligent authoring." Technology in medical writing will be increasingly adopted by the life sciences industry, and the future is not about the why, it is about the how. It is about how do you successfully implement it at scale. The industry is still stuck in a "pilotitis" mode, that is, operating on running one pilot to see if the technology really works, which is not surprising because it is such a highly regulated industry.

What are the areas in which use of the technologies might not be the best option and what barriers still exist in the industry?

Helle Gawrylewski: Aspects that require expert scientific knowledge and assessment may be able to be produced by automation but at this time still require human evaluation and judgment. Electronically translated text still needs human review because language nuances and cultural aspects are difficult to program, especially in many languages. A native speaker should always review and verify. Writers work in a global arena and should take this responsibility very seriously. For safety narratives, the medical assessment is also better written by a qualified medical writer. Machine written text can take on a repetitive quality and be interpreted as obviously machine written and not properly evaluated.

Nimita Limaye: Absolutely—addressing scientific and cultural nuances is critical. And I believe that writing is not just a science, it is an art. The sentience that a human can bring in can make all the difference, especially when it comes to developing lay summaries or building out informed consent forms. In addition, interpreting findings often requires looking across multiple data points, possibly in different reports. Algorithms may not be configured to do that. This is where the scientific thinking that a medical writer brings to the table counts.

What barriers to adoption exist in the industry?

Helle Gawrylewski: Structured authoring has been difficult to adopt because the application initially did not support Word documents and the formatting was an issue. Using structured authoring requires staff training and often an authoring tool does not integrate well with other older systems. It's easier for a new operation or initial public offering (IPO) to start with structured authoring than to have a large organization scrap all the old systems and replace them. Cost is definitely an issue but also technical competency of the staff. Even Word is not actually used to its full capacity! Document experts are often not writers, and many writers are not sophisticated technology experts! In the past writers have been reluctant to embrace automation because they think it will replace them. But the fact is that not everything they write is worthy of their full attention. So, offloading what can be offloaded allows full-time focus on the critically important document sections and elements.

There are some specific phases of research or types of research documentation for which automation seems to be more useful than for others.

In early phase 1 studies, much of the results are focused on data and assessments are straightforward, like blood levels for C_{max}, AUC, and such. Wearable devices that record results digitally are ideal in many types of studies in which tracking is important and in which some participants can be unreliable, such as in cardiac and diabetes studies.

Automation of patient diaries has always been a good use of automation, and now it's possible to use smartphones and audio recording to get quality real-time data.

Nimita Limaye: I think that the biggest barriers to adoption are change management and "pilotitis." Automation creates concerns with many medical writers. Will their roles be replaced? No, not really. They will actually move up the value chain. The grunt work will be taken care of by AI/ML. The medical writer will need to ensure that the data are represented in the right way, are being interpreted appropriately, and that the messaging is correct. It is important that the value of automation of medical writing is recognized. Secondly, implementing any technology requires investment, and returns come when the solution is implemented at scale. Hence, many times, companies do not see the returns after running a pilot, and then determine that this is not a good solution. That should not happen. Skill development is also important. Not everyone is tech-savvy, and the ability to navigate various tools requires training. Ensuring transparency and regulatory compliance will be critical.

What promising developments in automation exist in the near future as advances in AI and deep learning technologies continue to evolve?

Helle Gawrylewski: Access to efficient and useful information from large databases that are untapped and useless to regular human review, like Clintrial.gov, can have considerable impact. How many people can review and get value from all of the studies registered and reported there? The information is only as useful as we can accurately search and summarize it using AI and other newer methods of deep learning. I've seen it done and can say that it's exciting and not used nearly as much as it could be. The same applies to the EU portal that will contain not only CSRs but entire applications and IBs.

Workflows can be made efficient and accurate using automation and AI by an authoring system that reuses, and is connected with, all data elements linked for easy searching, correcting, and replication. Providing drug labels globally in all native languages that are accurate (and correspond with the master label), accurately translated, localized, and kept up to date in a master system for tracking and updating. I hope this will be more common than it is now. I think an AI and deep learning system to render research into plain language to make it accessible to the public could be a remarkable way to counter misinformation and shine a light on all the great scientific research that goes on but is inaccessible to most. It's said that the vaccines were developed so quickly that they can't be safe—wrong! The platforms were used for years before for other vaccine development, especially the standard ones used for Ebola and tuberculosis, but the public finds it hard to follow or understand what goes on in research. And we need to modernize regulatory processes and health authority reviews, continue to have applications and data digitally accessible and reviewable globally. If we are transparent and share, scientific data will help us make better decisions faster and promote not only cures but the prevention and avoidance of disease.

Nimita Limaye: One will see the increasing use of real-world data; data will be flowing in, fast and furious. It will be extremely challenging for medical writers to handle this scale and speed. This is where technology will play a valuable role. In addition, as global regulations keep evolving, dynamic document templates that embed this intelligence real-time will reshape the future of medical writing.

Author declaration and disclosures: *The author notes no commercial associations that may pose a conflict of interest in relation to this article.*

Author contact: *nlimaye@idc.com*

BIOSUMMARIES

Helle Gawrylewski has a MA from University of Pennsylvania, is a Woodrow Wilson Fellow, and is a former Senior Director in global regulatory and medical writing at Johnson & Johnson (J&J) (retired). Her experience in regulatory medical writing and global regulatory affairs spans more than 49 years in the pharmaceutical industry at Hoechst Roussel Pharma, Novo Nordisk, and Janssen research and development of J&J. During that time, Gawrylewski was directly involved with 55+ regulatory applications for marketing approval and in all aspects of product life-cycle development, ranging from early to full development and post-marketing medical affairs.

She managed and mentored staff from two to 125 and is a strong proponent and advocate of regulatory medical writing. She established linguistic services like translation and related global partnerships in medical and regulatory writing, leading outsourcing relationships in India and China and worked on the first team to submit a drug application electronically to the FDA. She led document management implementation and transparency activities internally while serving as a team lead at TransCelerate in the Clinical Trial Document Transparency group, later in PHUSE as a team member, and also on teams at Janssen that produced several European Medicines Agency Policy 0070 submissions of transparent clinical reports. In regulatory, she established global labeling outsourcing. Externally, Gawrylewski was the Pharmaceutical Research and Manufacturers of America representative in the ICH E3 Q&A working group that clarified standards for study reports, was a member of the CDISC Glossary Team and was the lead for 7 years, and was DIA MW community lead for and a core team member for 8 years. Gawrylewski is dedicated to cross-industry groups designing approaches to common problems in clinical trials, including clear goals/design, auditable conduct, subsequent clear reports, and transparent results in plain language and well-defined scientific terms shared in multiple languages. She is a member of the Multi-Regional Clinical Trials Plain Language Glossary effort, the PHUSE Transparency Term Harmonization Team, and contributed 2 chapters to the Regulatory Affairs Professionals Society's Regulatory Writing: an Overview. Experience shows that such work allows medical knowledge to advance and ultimately to make a difference in patients' lives.

Nimita Limaye, PhD, is a Vice President of Research with IDC Health Insights and leads Life Sciences Research and Development Strategy and Technology, providing research-based advisory and consulting services as well as market analysis on key topics related to the life sciences industry with a technology lens. She is an executive business leader with over 25 years of experience working in the pharmaceutical, contact research organization, and life sciences technology consulting industries. She is the past chair of the Society for Clinical Data Management board and is the current chair of the global DIA medical writing community. She has chaired several conferences, led industry roundtables, given keynotes, and has authored close to 100 publications and white papers. Limaye has led medical writing operations, managed strategic outsourced partnerships, and has conducted workshops on the outsourcing of medical writing.