Considerations for the use of artificial intelligence in the creation of lay summaries of clinical trial results

Kimbra Edwards[1]
On behalf of the working group members
[1] Health Communication Services, Center for Information & Study on Clinical Research Participation, Boston, Massschusetts, USA

doi: 10.56012vmag9372

Correspondence
Kimbra Edwards
kedwards@ciscrp.org

Abstract

The clinical research landscape is constantly evolving, as new regulations and innovations come together to help accelerate scientific discoveries and medical advances. A prominent example of this is the rapidly emerging technology of artificial intelligence (AI). Using AI to develop lay summaries (LS) of clinical trial results can enhance transparency and accessibility, while maximising efficiencies and facilitating scalability. This document is a product of collaboration between experts from over 15 organisations in the US and the EU, including industry, academia, and a patient-focused nonprofit. It aims to explore how AI can be responsibly applied to LS development. While aligning with current industry standards, this document provides several recommendations for AI implementation that highlight the necessity of human oversight and expertise. This joint effort between human and machine can help LS achieve high standards in accuracy, transparency, and compliance, while building public trust and empowering patients to make informed healthcare decisions.

Introduction

The landscape of pharmaceutical research is constantly evolving as emerging technologies reshape conventional practices. One such advancement is the use of artificial intelligence (AI) to support the development of lay summaries (LS) of clinical trial results. LS play a crucial role in increasing transparency and ensuring that trial results are accessible and understandable to patients, their caregivers, and the wider public. As AI technology rapidly evolves, it presents both considerable benefits while also introducing risks that must be thoughtfully managed in the context of LS development.

This document reflects the collaborative efforts of a diverse working group consisting of over 15 organisations from the US and EU, representing industry, academia, and a non-profit patient-focused organisation. The working group is composed of professionals with expertise in medical writing, technology, clinical operations, plain language, clinical trial transparency, and patient engagement. Together, they explored how AI can be responsibly applied to the creation of LS.

This document aligns with all broadly accepted industry standards, particularly the Good Lay Summary Practice (GLSP) guidance.

Background

Lay summaries are designed to make clinical research results more accessible to non-scientific audiences by translating complex medical information into plain language. AI can improve the efficiency of drafting LS by reducing manual effort (i.e., time and resources). Whereas the lack of sufficient human oversight can lead to inaccuracies or misinterpretations. For example, using data solely from sources like ClinicalTrials.gov may lack the proper context to appropriately develop an accurate and complete LS.

AI applications in health care are increasingly subject to oversight from regulatory authorities. At the time this document was authored, the US and EU are developing frameworks aimed at ensuring data privacy, accuracy, and ethics, such as the US Blueprint for an AI Bill of Rights,[1] the NIST AI Risk Management Framework,[2] the EU Artificial Intelligence Act,[3] the EU Ethics

Working group members

Sanjay Bagani, Vice President, Clinical Trials Transparency, Xogene
Behtash Bahador, Director, Health Literacy, Center for Information & Study on Clinical Research Participation (CISCRP)
Sudipta Chakraborty, Associate Director, Head of Health Literacy & Plain Language Center of Excellence, Biogen
Kimbra Edwards, Senior Director, Health Communication Services, CISCRP
Julia Farides-Mitchell, Director, Health Communication Services, CISCRP
Zack Fey, Manager, Medical Writing, CISCRP
Ken Getz, Founder, Board Chair, CISCRP; Director, Tufts CSDD
Julie Holtzople, Independent Consultant, Clinical Trial Transparency Thought Leader
Jennifer Menzer, Senior Expert Clinical Trial Transparency, Merck Healthcare KGaA, Darmstadt, Germany
Anne McClain, Transparency Operations Lead, UCB
Madison Reider, Senior Manager, R&D Patient Partnerships, CSL Behring
Malgorzata Sztolsztener, Director, Clinical Trial Transparency Patient Engagement, AstraZeneca
Woo Song, Co-founder, Xogene
Jamila Watkins, Clin-Ops, Patient, Site and Community Engagement Expert
Kristi Whiteside, Clinical Trial Transparency Expert

AI and LS terminology

> In this document, **AI** will be used to refer primarily to large language models (LLMs) that generate text,such as GPT, Gemini, Claude, and Llama.
> Additionally, **lay summaries (LS)** of clinical trial results are also known as lay language summaries (LLS) of clinical trial results, plain language summaries (PLS) of trial results, or trial results summaries (TRS).

Guidelines for Trustworthy AI,[4] and the FDA's Guidance on the use of AI in the development of drugs and biological products.[5] In addition, organisations like ICMJE, AMWA, EMWA, and ISMPP emphasised the need for transparency, accuracy, and human oversight in AI-generated medical writing. In turn, at the enterprise level, sponsor organisations are developing and deploying AI use cases and policy documents, including tools for drafting scientific and public- or patient-facing documents.

As AI regulatory frameworks, guidelines, and technology evolve, stakeholders must stay informed and adopt best practices to ensure high quality and compliant LS. At the time of this writing, there were limited guidelines on AI use in medical information, and none specifically addressing LS or other patient-facing clinical research information.

This document was initially drafted using AI to evaluate the feasibility of the outlined recommendations and considerations we have developed. Human experts reviewed and revised the content through multiple iterations, including feedback from a public comment period. The feedback received during this period was largely from research professionals, but also included some patients or members of the public. Both human review and AI were used to review and revise drafts for tone, spelling, and grammar.

Opportunities and risks

AI has the potential to enhance the efficiency of LS creation and promote health literacy by supporting broader dissemination of clinical trial results to patients and the public in a faster and cost-effective way. When used effectively, AI can streamline development, allowing for quicker delivery of clear, concise content. With reduced resource demands, research sponsors may be able to develop LS for more trials. However, without appropriate safeguards, AI-generated LS may contain inaccuracies. Potential issues include hallucinations (false, fabricated or misleading information that may arise when the AI model does not have adequate input and training), lack nuance (especially as it relates to scientific data), and insensitivity to tone and culture.

A hybrid approach, where AI supports drafting and experts ensure accuracy and appropriateness, can maximise benefits while minimising risks. This way, sponsors can maintain regulatory compliance while ensuring the public and patients receive timely, understandable, and accurate information to make effective decisions about their health.

Application and scope

This document supplements existing LS development practices, including the GLSP guidance, which remains the accepted industry standard for creating and delivering high quality LS. The GLSP has been adopted by the Clinical Trials Expert Group, a working group of the European Commission. It is not the intent of this document to replace general best practices for writing LS. Instead, it offers key considerations for incorporating AI into established LS workflows. These principles could also apply to other public- and patient-facing materials, such as lay language protocol synopses.

The recommended approach emphasises that responsible AI use is critical – it is to complement, not replace, human expertise. By combining thoughtful AI use with expert review, we aim to create clear and useful materials that enhance public and patient understanding of clinical research and promote accuracy, transparency and trust.

Considerations

Human involvement

Standalone use of AI for creating LS concerns clinical trial sponsors because AI lacks the nuanced understanding and contextual knowledge human experts provide in understanding complex clinical trial data. Without proper oversight, AI-generated LS may misrepresent these data or miss critical details leading to inaccurate summaries. This concern was observed in 2023 in a large-scale instance of publicly posted, AI-generated LS that lacked proper human oversight. These lay summaries were eventually removed from the public domain after significant concerns were raised regarding their accuracy. To better ensure accuracy and appropriate tone, AI should complement, not replace, human expertise and review, with professionals and members of the target readership reviewing and refining content.

Disclosure of AI use

Transparency regarding AI involvement in developing LS is essential for maintaining public trust and upholding ethical standards. Failure to disclose AI involvement can lead to skepticism, undermine confidence in the information, and damage the credibility of the author or organisation.

As AI capabilities continue to advance, open communication helps address misconceptions about the technology and build a more informed and trusting relationship between the public and the research community. To support this, clear disclosure of AI involvement, the extent of human oversight, compliance with regulations such as the EU AI Act, and acknowledgment of sponsor or patient community involvement are important considerations. See Appendix C for additional guidance and example disclosure statements.

Research sponsor involvement

Research sponsors are responsible for the study design, objectives, endpoints, and interpretation of results. Their input is vital for ensuring that LS accurately reflect trial findings and for precise interpretation of complex data. As public access to trial results increases, isolated creation of LS by external parties risks misinterpretation and loss of important context. While improved accessibility tools can promote equity in information dissemination, the absence of sponsor oversight has been demonstrated to lead to misinterpretation or omission of important details in the LS.

EU Artificial Intelligence Act disclosure guidance

> Per Chapter 4, Article 50, Paragraph 4 of the EU Artificial Intelligence Act:
> *Deployers of an AI system that generates or manipulates text which is published with the purpose of informing the public on matters of public interest shall disclose that the text has been artificially generated or manipulated.*

Misinformation and disinformation

Misinformation refers to unintentional errors, that can occur when AI misinterprets data or lacks the context to understand scientific concepts. Disinformation, on the other hand, is the deliberate distortion of facts with the intent to mislead. Either issue may arise if the AI systems being utilised are open-source or trained on public data without proper vetting. The opaque nature of AI decision-making compounds these risks.

Implicit bias and cultural sensitivity

Bias in training date or user prompts – whether intentional or unintentional – can lead to biased outputs. When AI models are trained on large datasets that may not fully reflect the diverse cultural backgrounds, the generated content can lack cultural awareness and sensitivity. AI can reproduce and even amplify those biases, resulting in skewed summaries that compromise the objectivity of information shared with patients and the public.

Promotional tone

LS should be written in a neutral, non-promotional tone. AI models are trained on large datasets, potentially including marketing content, which may result in the use of persuasive or overly positive language. This can bias the presentation of results, potentially misleading readers about the study's significance, benefits, or risks. In a clinical research context, maintaining a neutral, factual tone is essential to accurately convey findings and uphold public trust.

Rapid technological change

AI technologies evolve quickly, and using outdated models may lead to inaccuracies and inconsistencies in the generated content. This rapid pace of change may also make it challenging to keep AI tools aligned with the latest standards and best practices. This could increase the risk that LS may not meet current regulatory or quality expectations. With appropriate AI governance (see Appendix B for more details) this risk can be mitigated effectively.

Data privacy

Clinical study data sets contain sensitive personal health information about the participants. To ensure data privacy is maintained, all inputs used to create the LS should not include identifiable patient data. Aggregated data should be used, and organizations must ensure that AI models are not retaining sensitive information. Good data stewardship is required.

Recommendations for effective AI use in LS development

AI is a transformative tool that can enhance productivity in LS development. Examples of productivity include handling repetitive tasks like drafting, organising information, and simplifying technical language. It's important to ensure that all machine-generated outputs are reviewed by humans, who bring essential judgment in areas where AI may fall short. By using AI to support –not replace – human expertise, organisations can improve efficiency while ensuring LS remain accurate, appropriately written for their target audience, and aligned with regulatory standards.

Suggested additions to process flow

AI should be integrated at specific points in the existing LS development process, such as the best practices and overall process (as laid out in the GLSP) with clear roles for human review and approval (see Figure 1).

Key stakeholders and expertise

The effectiveness of AI in generating LS is contingent upon the expertise of the humans involved in its training, prompting, oversight, generation and revisions of LS. To ensure adherence to best practices and maintain quality and accountability, all reviewers and approvers recommended by the GLSP should retain their essential roles, skills, and qualifications in the LS process, even when AI tools are integrated. While standard operating procedures and resourcing at organisations may vary, stakeholders possessing the following additional AI knowledge and experience may play critical roles at various stages:

- **AI training and development experts**: AI development experience is required to design and calibrate the AI systems to properly train the system on relevant inputs and datasets.
- **Health literacy specialists**: Expertise in health literacy and plain language writing should be leveraged to help train the AI on simplifying complex medical language into terms that are understandable, including guiding AI on which terminologies, explanations, and formatting best align with the needs of the reader.
- **Legal and compliance teams**: To ensure that AI systems use data safely and in accordance with approved AI and/or data use laws and policies (such as GDPR or HIPAA), appropriate expertise should be incorporated into training, building the appropriate framework. Data privacy monitoring can be achieved through standard LS review procedures.
- **LS and medical writers**: Once AI generates a draft, medical writers will need to ensure the AI's interpretation of clinical results are factual and that no critical scientific nuances are absent in the LS. This will be different than what they have traditionally done in authoring this information for the LS.

Additional considerations

It is important to recognise that while LLMs are capable of generating human-like text, they still have limitations to be managed. This section outlines several additional considerations when implementing AI for LS.

- **Templates and glossaries**: Standardised templates can help ensure consistency and compliance with regulatory requirements. AI should be trained to work within these templates while allowing for necessary flexibility such as study design and/or different therapeutic areas. AI should also be trained to use a glossary for preferred terminology within a particular document or set of documents and previously completed summaries.
- **Data inputs**: The quality and comprehensiveness of data inputs are crucial for generating accurate and relevant LS. Reducing the risk of

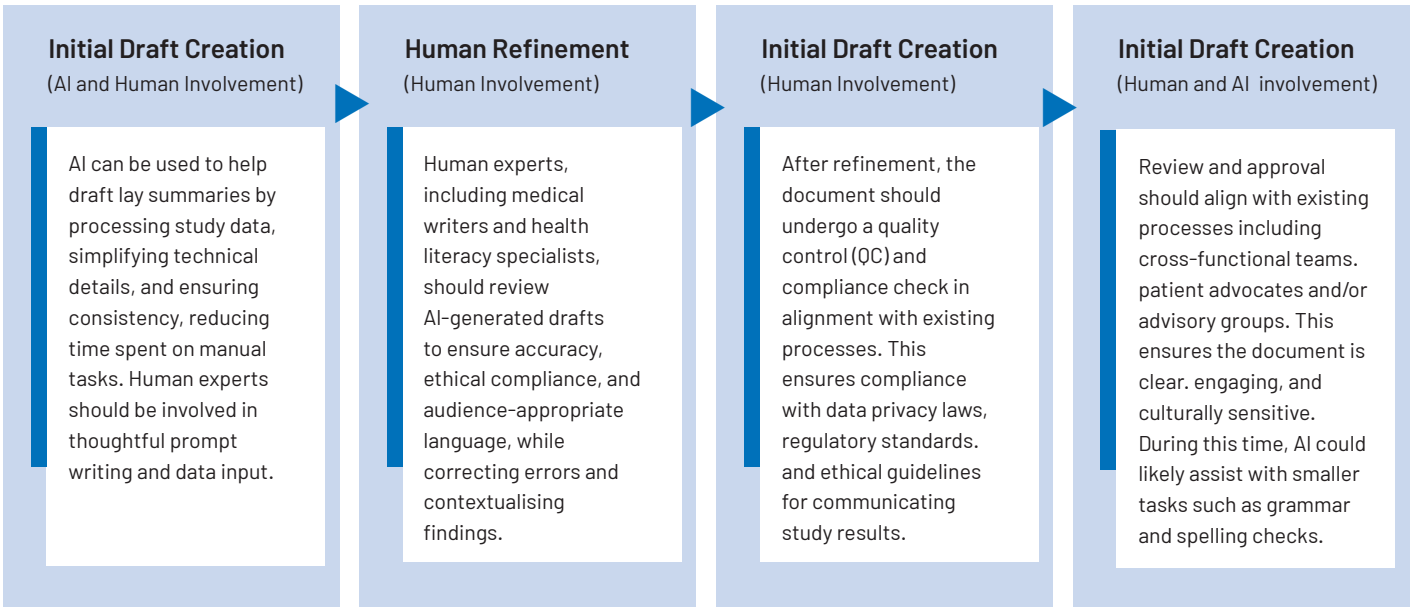

Figure 1. Suggested additions to process flow

AI hallucinations or overconfidence helps prevent seemingly legitimate responses that may omit critical information or draw incorrect conclusions. It would be beneficial for AI models to include references from source documents from which data and information are being pulled. Key data sources may include:

- Aggregate tables, figures, and listings (TFLs)
- Clinical study protocols (CSP)
- Clinical study reports (CSR)
- Informed consent forms (ICF)
- Lay protocol synopsis
- Other nonpromotional public or patient-facing documents
- Glossaries of medical terms and plain language equivalents

- **Prompt engineering:** A critical component of using AI effectively is prompt engineering, which guides the AI in creating accurate, understandable, and public- and patient-appropriate content. For each LS document to be drafted multiple and sequential prompts should be provided to the AI for drafting individual sections and for clear context setting. Specific instructions on tone and style, and guidelines for simplifying complex concepts should be provided. These prompts help the AI create the right tone, ensure consistency with approved medical terminology, and address potential biases. By including reminders to provide necessary context and caveats, prompt engineering can help ensure that AI-generated content is both informative and patient-friendly. Please see Appendix A for components of good prompts and example prompts.
- **Governance**: Robust AI governance is essential for overseeing any new system including an AI system. Implementing AI is an iterative process that requires initial testing and continuous improvement. Please see Appendix B for additional considerations.
- **Advanced AI architectures**: Leveraging AI most effectively may require more advanced architecture, such as AI agent networks. Agent networks employ multiple AI agents, each with a specialised role such as a medical fact-checker, readability optimiser, and bias and sensitivity detector. Orchestrator agents can also be integrated into the architecture to coordinate the work of specialised agents, like a project manager, while humans continue to provide expert oversight and intervention at key points.

Organisations can harness the potential of AI to enhance their LS processes through carefully addressing both opportunities and risks outlined in this document, and through continuous learning. Regular monitoring and updates to processes and AI models with the latest medical and regulatory information will likely be essential to mitigate associated risks and maintain the highest standards of accuracy, clarity, and ethical LS practice.

Conclusion

Incorporating AI into LS development presents both opportunities and risks, underscoring the need for thorough planning and careful implementation. While AI can improve efficiency and reach, its output must be guided by human expertise to ensure accuracy, sensitivity, and compliance. Successful implementation will be an ongoing process that requires continuous monitoring, evaluation, and refinement. Ultimately, integrating AI into LS development necessitates balancing innovation with oversight, ensuring each summary meets the highest standards of quality, accuracy, and transparency, better ensuring trust and clarity for patients and the public.

References

1. The White House, Office of Science and Technology Policy. Blueprint for an AI Bill of Rights: Making automated systems work for the American people. Washington, DC: The White House; 2022 [cited 2025 Apr 10]. Available from: https://bidenwhitehouse.archives.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf
2. National Institute of Standards and Technology. Artificial Intelligence Risk Management Framework (AI RMF 1.0). Gaithersburg (MD): U.S. Department of Commerce; 2023 [cited 2025 Apr 10]. Available from: https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf

3. European Union. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) 2018/858, (EU) 2019/2144 and (EU) 2023/1230 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 [Internet]. Brussels: EUR-Lex; 2024 [cited 2025 Apr 10]. Available from: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024R1689
4. European Commission, High-Level Expert Group on Artificial Intelligence. Ethics guidelines for trustworthy AI. Brussels: European Commission; 2019 [cited 2025 Apr 10]. Available from: https://ec.europa.eu/futurium/en/ai-alliance-consultation.1.html
5. FDA. Considerations for the use of artificial intelligence to support regulatory decision-making for drug and biological products. Silver Spring, MD: FDA; 2025 [cited 2025 Apr 10]. Available from: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/considerations-use-artificial-intelligence-support-regulatory-decision-making-drug-and-biological

Author information

Kimbra Edwards, PhD, is the Senior Director of Health Communication Services at the Center for Information and Study on Clinical Research Information (CISCRP), where she oversees the creation of easy-to-understand trial resources for patients, participants, and the public.

Appendix A – Prompt engineering considerations

Components of good prompts:

- Clear context setting (e.g., "You are writing a lay summary for a clinical trial on [condition] for people with a 6th-grade reading level.")
- Specific instructions on tone and style (e.g., "Use a compassionate and encouraging tone while maintaining factual accuracy.")
- Guidelines for simplifying complex concepts (e.g., "Explain [medical term] in simple language a non-expert can understand.")
- Reminders to include necessary context and caveats (e.g., "Ensure to mention that these results may not apply to all patients and individual responses may vary.")
- Use of sequential prompts for refinement can help improve the quality of a draft.

Example prompts:

- "Please create a lay summary of clinical trial results for a new diabetes medication. Your audience is the general public, including patients with type 2 diabetes, who have a 6th-grade reading level. Use a compassionate and encouraging tone while maintaining factual accuracy. Simplify complex medical terms but include them in parentheses after the simplified explanation. Ensure you mention the study's limitations and that results may not apply to all patients. Structure the summary with understandable headings and bullet points for easy readability."
- "Please write a 3-paragraph explanation for why this trial: [trial name and NCT number from publicly available website] is being done. In the first paragraph please explain the condition, in the second paragraph please explain the study drug and why it is being developed, and in the third paragraph please discuss the trial design and restate the hypothesis for the final sentence. Please write the entire explanation at a 12 year old reading level."
- "You are tasked with creating a lay summary of clinical trial results for a new diabetes medication. Your audience is the general public, including patients with type 2 diabetes, who have a 6th-grade reading level.
- Here are the clinical trial results you will be summarising: [insert documentation if within LLM capabilities/applicable].

Follow these guidelines to create your summary:

1. Use a compassionate and encouraging tone throughout the summary. Be warm and supportive but maintain factual accuracy.
2. Write at a 6th-grade reading level. Use simple words and short sentences. Avoid jargon or complex medical terminology.
3. Structure your summary with the following headings:
   - What was the study about?
   - What did the study find?
   - What does this mean for me?
   - What are the next steps?
4. Under each heading, use bullet points to present information clearly and concisely.
5. When introducing medical terms or concepts, first provide a simple explanation, then include the technical term in parentheses. For example: "sugar in the blood (glucose)".
6. Mention the study's limitations and clearly state that the results may not apply to all patients.
7. Begin your summary with a brief overview of the study's purpose (2–3 sentences).

Write your complete summary inside <summary> tags. Ensure that your summary is factually accurate based on the provided clinical trial results, while being easy to understand for the target audience."

Appendix B – Considerations for AI governance

Effective governance is crucial when implementing AI for plain language summaries. A well-structured governance framework ensures that the use of AI aligns with organisational goals, regulatory requirements, and ethical standards. Key components of governance should include:

- **Internal collaboration & standards development/implementation**
  - Establish a cross-functional team including medical writers, statisticians, legal experts, patient advocates, and AI specialists.
  - Develop clear guidelines and standard operating procedures (SOPs) for AI use in patient communications.
  - Implement a review and approval process involving subject matter experts to validate AI-generated content.
  - Create a feedback loop to continuously improve AI performance based on human expert input.

- **Initial testing**
  - Develop a comprehensive test suite covering various scenarios, e.g., study phase, design, endpoints, safety data sets, patient populations
  - Conduct A/B testing comparing AI-generated content with human-written content for patient preference and understanding
  - Implement a feedback loop incorporating input from patients, healthcare providers, and subject matter experts
  - Regularly update and retrain AI models based on new data, feedback, and evolving best practices
  - Testing process example:

1. Generate initial content using AI
2. Review by humans for accuracy, readability, and health literacy levels using validated tools
3. Incorporate public and patient involvement for feedback on understandability and relevance
4. Iterate based on feedback, making necessary adjustments to prompts or AI models
5. Repeat steps 1–5 until satisfactory results are achieved
6. Implement in a limited rollout and monitor performance

Scale implementation based on successful performance metrics

- **Ongoing monitoring given AI's continuous learning**
  - Implement a phased rollout, starting with low-risk applications and gradually expanding to more complex tasks.
  - Establish key performance indicators (KPIs) to measure the accuracy, readability, and effectiveness of AI-generated communications.
  - Conduct regular audits to assess AI performance.
  - Implement a system for ongoing monitoring of AI outputs, including random sampling and human expert review.
  - Develop protocols for addressing and correcting any errors or biases identified in AI-generated content.
  - Stay informed about advancements in AI technology and update systems accordingly to maintain state-of-the-art performance.

- **Regulatory compliance**
  - Ensure compliance with relevant regulations, such as the EU AI Act, GDPR, and FDA guidelines.
  - Maintain detailed documentation of AI training data, algorithms, and decision-making processes for regulatory audits.
  - Establish a process for staying updated on evolving regulations and adjusting AI systems and governance practices accordingly.

- **Ethical considerations**
  - Develop an ethical framework for AI use in patient communications, addressing issues such as bias, privacy, and transparency.
  - Implement safeguards to protect patient data and ensure confidentiality throughout the AI-assisted communication process.
  - Regularly assess the ethical implications of AI use and make necessary adjustments to maintain alignment with organisational values, industry best practices, and societal expectations.

- **Training and education**
  - Provide comprehensive training for staff involved in AI-assisted patient communication processes.
  - Develop resources to help team members understand AI capabilities, limitations, and best practices for collaboration between humans and AI systems.

- **Continuous improvement**
  - Establish a process for collecting and analysing feedback from patients, healthcare providers, and other stakeholders on AI-generated communications.
  - Use insights gained from feedback and performance monitoring to refine AI models and improve the quality of patient communications over time.

Appendix C – Considerations for AI disclosure

Transparency regarding the use of generative AI in creating patient communications is essential for maintaining trust, ethical standards, and regulatory compliance. Proper disclosure practices should address the following aspects:

- Where and when should the use of AI be disclosed and to what extent
  - Include a clear statement about AI involvement in the creation of the document, typically in the introduction or a dedicated section.
  - Disclose the extent of AI use, such as whether it was used for initial drafting, language simplification, or spelling/ grammar checking.
  - Consider including a brief explanation of how AI was used in conjunction with human expertise to ensure accuracy and relevance.
  - Make the disclosure easily understandable for the target audience, avoiding technical jargon.

- AI regulation compliance
  - Ensure that disclosure practices align with the requirements of the EU AI Act or similar, applicable regulations.
  - Provide information on the AI system's purpose, capabilities, and limitations as required by applicable laws.
  - Include contact information for inquiries about the AI system or its outputs.

- Disclosure of sponsor or other human involvement:
  - Clearly state the level of involvement of the study sponsor and medical experts in reviewing and approving the LS.
  - Acknowledge any public or patient community involvement in the development or review of the LS.
  - If there was limited or no human involvement, this should also be disclosed transparently.

Example disclosure statements to include in LS:

- AI involvement disclosure
  - "This summary was initially drafted using artificial intelligence (AI) technology. After the first draft was created, it was reviewed, revised, and approved by qualified medical professionals to ensure accuracy, clarity, and relevance."

- Extent of AI use
  - "Artificial intelligence was used to assist in simplifying complex medical language and organising information in this summary. All content has been verified and approved by the study team and patient representatives."

- Sponsor involvement
  - "The study sponsor, [sponsor name], has reviewed this AI-assisted summary to ensure its accuracy and alignment with the clinical trial results."

- Public and patient involvement
  - "Members of the public, patients, and patient advocates were also involved in the review of this summary to help ensure it is understandable and relevant."

- AI regulation compliance
  - "This document was created with the assistance of an AI system developed by [company name]. The system is designed to simplify medical language and organise information for LS. For more information about the AI system used, please contact [contact information]."

Appendix D – Example of advanced AI architecture for LS creation

Advanced AI systems for creating LS benefit from specialised agentic architectures that divide complex tasks among multiple AI components working in coordination. This approach mirrors team-based document creation in traditional settings but offers enhanced consistency, scalability, and traceability. In a sense, this approach is modeling human excellence.

Key components of an agentic architecture

Planning and creation agents

- **Strategy planning agent**:
  Analyses source documents and develops structural approach
  - Maps information complexity and creates audience-appropriate templates
  - Sets measurable objectives (reading level, length, key messages)
- **Initial drafter agent**:
  Transforms clinical documents into first-draft summaries
  - Structures information logically while maintaining appropriate detail balance
  - Adheres to target reading level parameters
- **Medical accuracy checker agent**:
  Verifies factual correctness
  - Cross-references claims against source documentation
  - Flags statistical information requiring expert verification
- **Readability optimiser agent**:
  Refines language for target audience
  - Adjusts text using readability metrics
  - Suggests simpler terminology while preserving meaning
- **Bias and sensitivity reviewer agent**: Ensure inclusive content
  - Identifies potentially exclusionary or stereotyping language
  - Checks for balanced representation and culturally sensitive explanations

Coordination and feedback

- **Orchestrator agent**:
  Manages workflow and integration
  - Routes content between specialised agents
  - Resolves conflicts and maintains document integrity
  - Identifies areas requiring human intervention
- **Feedback integration agent**:
  Processes human expert input
  - Categorises and prioritises feedback
  - Updates agent parameters based on feedback patterns

Human integration

- **Human expert touchpoints**:
  Strategic oversight at key junctures
  - Review of planning outputs and initial parameters
  - Evaluation of flagged uncertainties requiring domain expertise
  - Provision of structured feedback and final approval

Implementation workflow

1. **Planning**: Strategy agent analyses source documents and establishes approach.
2. **First draft**: Initial drafter produces structured summary based on planning.
3. **Multi-agent review**: Medical accuracy, readability, and bias agents evaluate draft.
4. **Integration**: Orchestrator consolidates agent inputs into revised draft.
5. **Human feedback**: Experts review and provide structured feedback.
6. **Refinement**: Agents implement changes based on feedback.
7. **Iteration**: Steps 5-6 repeat as needed until quality thresholds are met.
8. **System Learning**: Feedback patterns update agent parameters for future projects.
9. **Approval**: Human experts provide final sign-off with complete process documentation.

Benefits

- **Specialisation**: Optimised agents for specific tasks
- **Traceability**: Clear documentation of decisions
- **Adaptability**: System learns from expert feedback
- **Scalability**: Consistent approach across document types

Testing and monitoring of advanced AI Architecture systems is critical and should be implemented according to the guidelines outlined in Appendix B.

Appendix E – Helpful tools and resources

Leverage existing tools and resources and develop additional, use-specific comprehensive resources to guide the development and use of AI for LS creation. The following tools, resources, and topics should be considered.

A. Quality control checklists

- Verification of medical facts and statistics against source documents (e.g., clinical study reports, published literature)
- Consistency checks, inter- and intra-document, with approved messaging and terminology
- Assessment of readability and health literacy levels
- Evaluation of cultural sensitivity and inclusivity
- Identification of potential biases or misleading statements
- Identification of oversimplified or illogical statements
- Compliance with regulatory requirements and internal guidelines

B. AI implementation, evaluation and benchmarking tools

- Quality assessment frameworks for measuring accuracy and readability
- If using AI agents, developer tools to help understand agentic decision-making processes
- Performance benchmarking tools to compare AI outputs against human-generated content
- Annotation tools for providing feedback on AI-generated content

C. Data privacy safeguards

- Data anonymisation and de-identification tools
- Secure file transfer protocols for sensitive information
- Access control systems to limit data exposure
- Encryption tools for data at rest and in transit
- Privacy impact assessment templates

D. Collaborative platforms

- Implement secure platforms for collaboration between AI systems and human experts
- Version control systems to track changes and approvals
- Annotation tools for providing feedback on AI-generated content
- Project management software to coordinate review and approval processes

E. Training resources for staff involved in using AI

- Develop comprehensive training materials for staff involved in AI-assisted LS creation
- E-learning modules on AI capabilities and limitations
- Good practices for human-AI collaboration
- Regular workshops and webinars on emerging AI technologies and ethical considerations

F. Additional resources

- Good Lay Summary Practice Guidance (GLSP)
- International Society for Medical Publication Professionals (ISMPP) position statement and call to action on artificial intelligence
- EMA artificial intelligence workplan
- Four principles for safe and responsible use of LLMs (EMA)
- Guiding principles on the use of large language models in regulatory science and for medicines regulatory activities (EMA)
- European Union (EU) Artificial Intelligence Act

List of abbreviations in this article

AI	artificial intelligence
AMWA	American Medical Writers Association
CSP	clinical study protocol
CSR	clinical study report
CTEG	Clinical Trials Expert Group
EMWA	European Medical Writers Association
EU	European Union
FDA	Federal Drug Association
GDPR	General Data Protection Regulation
GLSP	Good Lay Summary Practice
GPT	Generative Pre-trained Transformer
HIPAA	Health Insurance Portability and Accountability Act
ICF	informed consent form
ICMJE	International Committee of Medical Journal Editors
ISMPP	International Society for Medical Publication Professionals
LLM	large language model
LLS	lay language summary
LS	lay summary
NIST	National Institute of Standards and Technology
PLS	plain language summary
TFL	tables, figures, listings
TRS	trial results summary
US	United States