Transforming Global Value Dossier (GVD) Drafting: Creation With a Generative Artificial Intelligence (GenAI)-Driven Coauthoring Accelerator

Gofman L[1], Meyerink J[2], Sharma S[3]

[1]ZS Associates, Princeton, NJ, USA, [2]ZS Associates, Los Angeles, CA, USA, and [3]ZS Associates, New Delhi, India.

Objectives

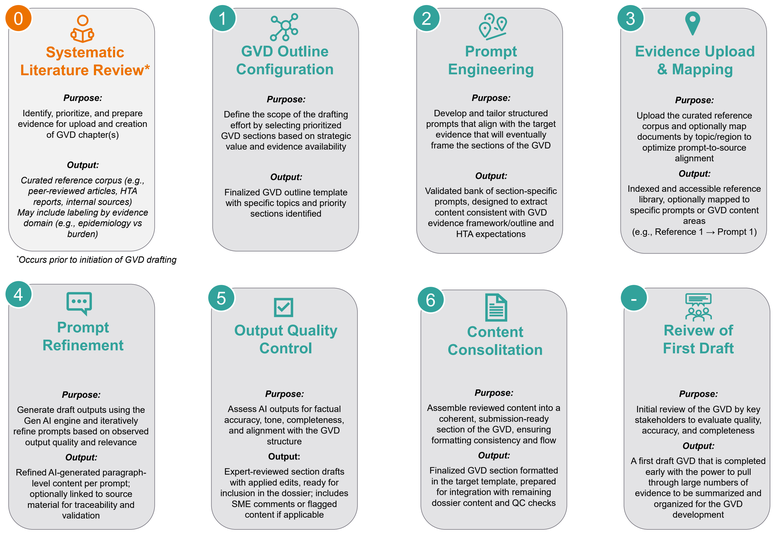
- As global value dossiers (GVDs) are critical in informing health technology assessments (HTAs) and payer decisions, efficient and accurate content creation is paramount[1,2]
- Development of GVDs is a labor-intensive, repetitive, and time-consuming process, requiring the synthesis of large volumes of evidence into well-structured narratives aligned with stakeholder expectations[1,2]
- Manual drafting introduces risk of inconsistencies, version control issues, and delayed timelines
- This study evaluated the generalizability and efficiency of a Gen AI-based Coauthoring Accelerator in producing high-quality Disease Overview content for GVDs (**Figure 1**)
- Specifically, we assessed content accuracy, completeness, usability, and time savings relative to manual drafting

Methods

Study Design

- A curated Disease Overview template was applied to 75+ reference documents covering breast cancer
- A retrieval-augmented generation (RAG) framework guided extraction from epidemiology, pathophysiology, risk factors, clinical presentation, diagnosis, and unmet needs data
- Prompts were engineered to prioritize scientific accuracy and were fine-tuned to reflect preferred GVD section structure
- Gen AI-generated outputs were reviewed by three subject matter experts (SMEs) with ≥5 years of experience in HEOR/medical writing
- **Key evaluation metrics included accuracy** of extracted content, **completeness** of disease-specific information, **time** to first draft, and **user satisfaction** on clarity and utility (**Table 1**)

Figure 1. Illustrative Gen AI-Driven GVD Process



Results

- Figure 2 demonstrates the conceptual flow outlining stages in the GenAI tool, from outline creation to prompt creation, leading to mapping of the prompt with the references, and then to a consolidated output (**Figure 2**)
- Predetermined evaluation metrics were defined through quality, accuracy, and performance to evaluate the GenAI-GVD tool (**Table 1**)
- The tool achieved a high extraction accuracy (~85%) and strong completeness scores (~92%) across key disease-related elements, including epidemiology, pathophysiology, and risk factors
- Draft generation time was reduced by approximately 60% relative to traditional methods, substantially accelerating the time to first GVD draft
- Reviewers reported high satisfaction (~70%), citing increased efficiency and the ability to focus efforts on strategic refinement rather than initial content creation (**Figure 3**)

Figure 2. Proof of Concept Outline and Prompt Retrieval Model

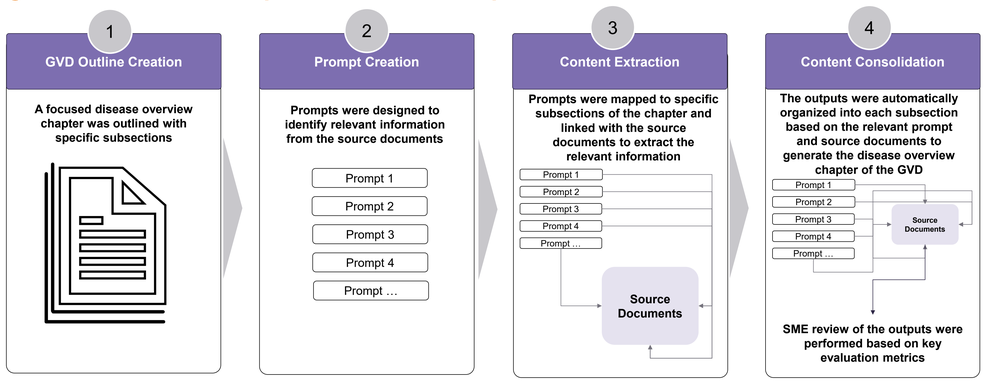
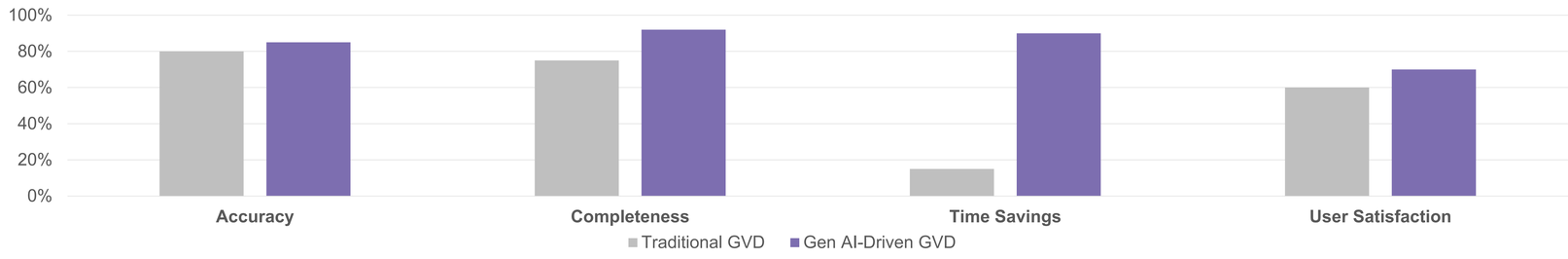


Table 1. Definitions of Key Evaluation Metrics

Evaluation Metric		Definition
Completeness	Comprehensiveness	Content covers all necessary aspects comprehensively.
	Coherence and Readability	Content is clear, well-structured, and easy to read, requiring few to no edits.
	Content Relevance	Uses inputs accurately in the appropriate context and summarizes content effectively with minimal edits needed.
Accuracy	Hallucination Mitigation	Content is cross-referenced via multiple reference sources to further ensure content accuracy, and user is notified when reference data are not present to support content creation.
	Contextual Understanding	Shows a deep understanding of context, producing content that fits seamlessly with minimal adjustments.
	Integration of Multiple Sources	Seamlessly integrates multiple reference sources into a cohesive and well-rounded output.
User Satisfaction	Efficiency in Content Generation	Generated content meets initial requirements and quality standards with minimal manual adjustments needed.
	Ease of Use and Level of Engagement	Tool is deemed accessible and easy to use by HEOR personnel, ensuring it does not add additional burden to the GVD drafting process.
	Scalability	Tool can handle large volumes and multiple GVD chapters efficiently.

Figure 3. Traditional GVD vs. Gen AI-Driven GVD Evaluation Comparison



Conclusions

- This study demonstrated strong efficiency, accuracy, and completeness in generating one chapter of a GVD by using our Gen AI-driven tool; leading to time saving and quality content with promising potential to streamline GVD development
- SME review is critical to ensure the generated outputs are both accurate and comprehensive
- With continued refinement and intelligent prompt optimization, we are providing efficiencies to scale Gen AI–enabled support across all chapters of the GVD—delivering consistent, high-quality drafts that are accurate, reproducible, and aligned with the evolving expectations of global HTA bodies

References

1. Liberti, L., & Wang, T. (2021). The regulatory-HTA decision-making interface: What the medical writer should know. Medical Writing, 30(3), 50-55.
2. Botwright, S., Sittimart, M., Chavarina, K. K., Bayani, D. B. S., Merlin, T., Surgey, G., & Teerawattananon, Y. (2025). Good practices for health technology assessment guideline development: a report of the Health Technology Assessment International, HTAsiaLink, and ISPOR Special Task Force. *International Journal of Technology Assessment in Health Care,* 40(1), e74.