

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

#### **Gofman L<sup>1</sup>, Meyerink J<sup>2</sup>, Sharma S<sup>3</sup>**

<sup>1</sup>ZS Associates, Princeton, NJ, USA, <sup>2</sup>ZS Associates, Los Angeles, CA, USA, and <sup>3</sup>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<sup>1,2</sup>
- Development of GVDs is a labor-intensive, repetitive, and timeconsuming process, requiring the synthesis of large volumes of evidence into well-structured narratives aligned with stakeholder expectations<sup>1,2</sup>
- Manual drafting introduces risk of inconsistencies, version control issues, and delayed timelines
- This study evaluated the generalizability and efficiency of a Gen AIbased 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 Al-generated outputs were reviewed by three subject matter experts (SMEs) with  $\geq$ 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 Al-Driven GVD Process   |   |  |  |
|--|---|--|--|
| 0 0  | 1 %   | 2 20   |  |
|  | 000   |  |  |
| Systematic   | GVD Outline   | Prompt   |  |
| Literature Review*   | Configuration   | Engineering  |  |
| Purpose:   | Purpose:  | Purpose:   |  |
| Identify, prioritize, and prepare<br>evidence for upload and creation<br>of GVD chapter(s)   | Define the scope of the drafting<br>effort by selecting prioritized<br>GVD sections based on strategic<br>value and evidence availability                 | Develop and tailor structured<br>prompts that align with the target<br>evidence that will eventually<br>frame the sections of the GVD  |  |
| Output:  |   | Output   |  |
| Curated reference corpus (e.g., peer-reviewed articles, HTA  | Output:   | <i>Output:</i><br>Validated bank of section-specific   |  |
| reports, internal sources)   | Finalized GVD outline template<br>with specific topics and priority   | prompts, designed to extract   |  |
| May include labeling by evidence   | sections identified   | content consistent with GVD  |  |
| domain (e.g., epidemiology vs<br>burden)   |   | evidence framework/outline and<br>HTA expectations   |  |
| Prompt   | 5<br>Output Quality   | 6 E<br>Content   |  |
| Refinement   | Control   | Consolitation  |  |
| <i>Purpose:</i><br>Generate draft outputs using the<br>Gen AI engine and iteratively   | <i>Purpose:</i><br>Assess AI outputs for factual<br>accuracy, tone, completeness,   | <i>Purpose:</i><br>Assemble reviewed content into a<br>coherent, submission-ready  |  |
| efine prompts based on observed  | and alignment with the GVD  | section of the GVD, ensuring<br>formatting consistency and flow  |  |
| output quality and relevance<br><i>Output:</i>   | structure<br>Output:  | Output:  |  |
| Refined AI-generated paragraph-<br>level content per prompt;<br>optionally linked to source<br>material for traceability and<br>validation | Expert-reviewed section drafts<br>with applied edits, ready for<br>inclusion in the dossier; includes<br>SME comments or flagged<br>content if applicable | Finalized GVD section formatted<br>in the target template, prepared<br>for integration with remaining<br>dossier content and QC checks |  |



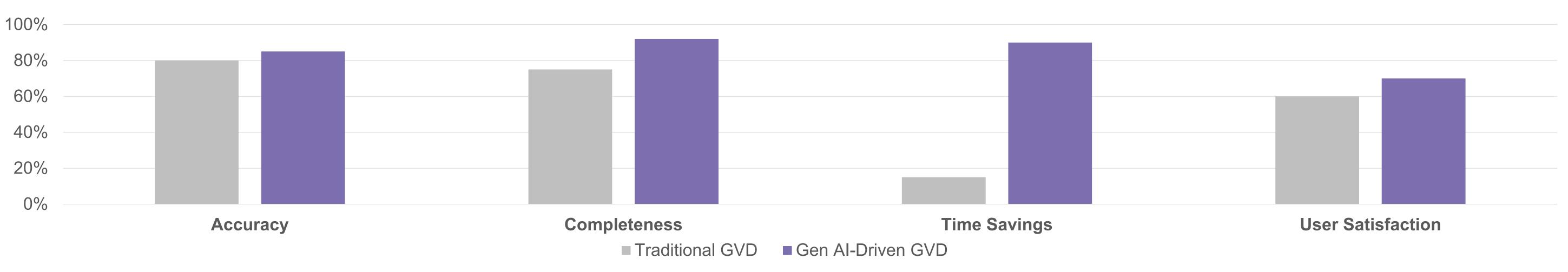
### Results

- epidemiology, pathophysiology, and risk factors
- initial content creation (Figure 3)

#### Figure 2. Proof of Concept Outline and Prompt Retrieval Model

| 1<br>GVD Outline Creation | Pr       | - 2<br>rompt Creatic  |
|---------------------------|----------|---|
| <text></text>             | identify | e source docu<br>Prompt 1<br>Prompt 2<br>Prompt 3<br>Prompt 4<br>Prompt |

#### Figure 3. Traditional GVD vs. Gen Al-Driven GVD Evaluation Comparison



### Conclusions

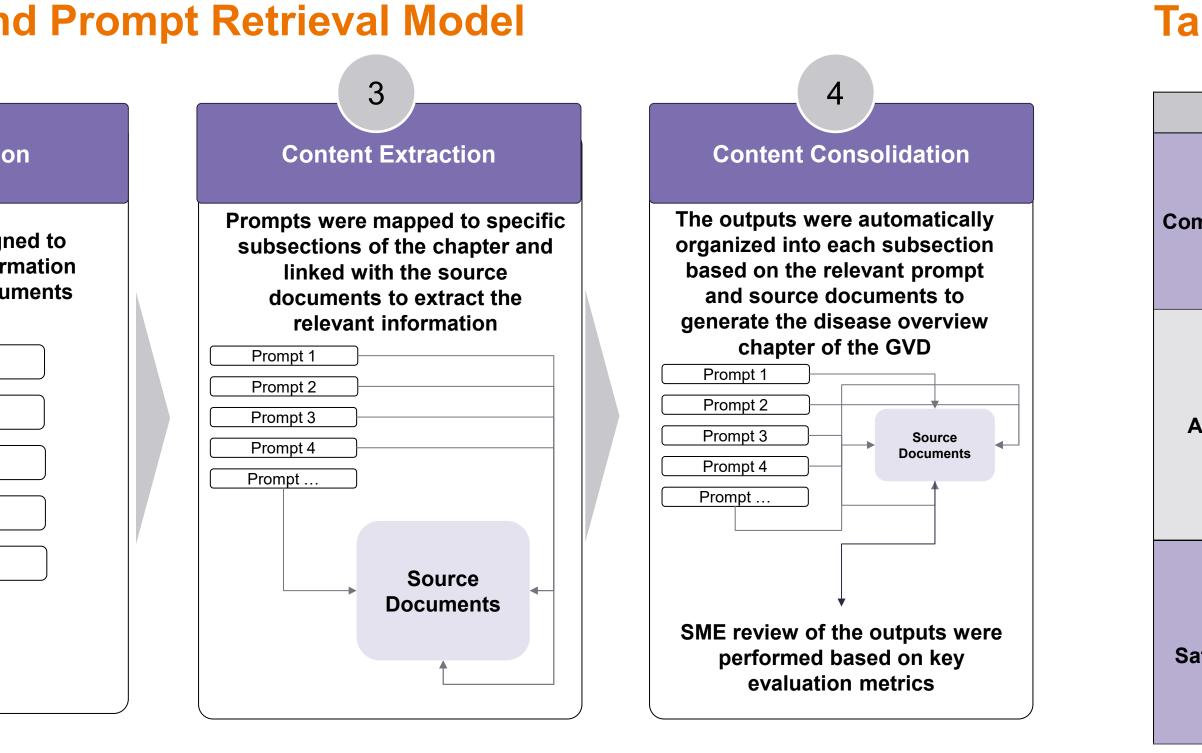
#### References

HTAsiaLink, and ISPOR Special Task Force. International Journal of Technology Assessment in Health Care, 40(1), e74.

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

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



• 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 Al-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

# **HTA 74**

#### **Table 1. Definitions of Key Evaluation Metrics**

| Evaluation Metric   |   | Definition   |  |
|---------------------|---|--|--|
| npleteness          | Comprehensiveness                         | Content covers all necessary aspects comprehensively.  |  |
|                     | Coherence and<br>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<br>Mitigation               | Content is cross-referenced via multiple reference sources to further<br>ensure content accuracy, and user is notified when reference data are<br>not present to support content creation. |  |
|                     | Contextual<br>Understanding               | Shows a deep understanding of context, producing content that fits seamlessly with minimal adjustments.  |  |
|                     | Integration of Multiple<br>Sources        | Seamlessly integrates multiple reference sources into a cohesive and well-rounded output.  |  |
| User<br>atisfaction | Efficiency in Content<br>Generation       | Generated content meets initial requirements and quality standards with minimal manual adjustments needed.   |  |
|                     | Ease of Use and<br>Level of<br>Engagement | Tool is deemed accessible and easy to use by HEOR personnel,<br>ensuring it does not add additional burden to the GVD drafting<br>process.   |  |
|                     | Scalability                               | Tool can handle large volumes and multiple GVD chapters efficiently.   |  |