Integrating stakeholder insights into thematic analysis: Enhancing efficiency and supporting HEOR

Cochrane J[1], Riley D[1], Yakob L[1], Heron L[1]
[1]Adelphi Values PROVE, Bollington, Cheshire SK10 5JB, United Kingdom

ADELPHI VALUES
PROVE Expertise in Access and Value Evidence Outcomes
SA56

Introduction

- Engaging with stakeholders and incorporating stakeholder insights is fundamental to ensuring that Health Economic Outcomes Research (HEOR) addresses the real-world needs, values, and priorities of those affected by healthcare decisions.
- Traditional methods of integrating stakeholder insights in HEOR involve the use of established qualitative methodology, including Delphi panels, interviews and focus groups.[1,2]
- Through these methods, extensive data sets can be obtained that can be subsequently thematically analysed to capturing the depth, context and nuances often missed by quantitative methods.[1,3]
- Thematic analysis, however, can often be labour-intensive, time-consuming, and may be subject to human biases through the subjectivity of interpretation or contextual differences.[4-6]
- Manual qualitative analysis typically involves multiple rounds of coding, theme identification, and synthesis, often requiring extensive coordination among research teams.[4] This process can take weeks or even months, limiting the frequency and depth with which stakeholder perspectives are incorporated into HEOR studies.[4,7]
- Artificial intelligence (AI) has emerged as a powerful tool to process, analyse, and generate insights from complex datasets across diverse fields, with potential to streamline thematic analysis of qualitative data in stakeholder research for HEOR.
- Large language model (LLM) AI tools, including ChatGPT and Claude, are able to rapidly analyse and process datasets of stakeholder data with minimal user involvement.[8]
- Therefore, they provide an opportunity to reduce turnaround times, support real-time analysis, and help standardise qualitative research processes, making it feasible to integrate stakeholder insights more systematically into HEOR.
- However, the analytical accuracy of these tools for qualitative analysis is currently unclear.
- We explored how LLMs can be implemented to improve the efficiency of thematic analyses of stakeholder data, whilst maintaining analytical accuracy, in the specific context of HEOR. We also examined the role of expert researchers in guiding and interpreting AI-generated outputs to ensure contextual depth.

Methods

- Six interview transcripts, from a diverse range of healthcare professionals (including paediatricians, haematologists, psychologists, A&E staff, and general practitioners [GPs]), focusing on unmet needs within a haematological disease area, were thematically analysed using ChatGPT-4.1 as part of an internal research initiative.
- All transcripts provided to the LLM were labelled to clearly identify the profession of the interviewee. However, all other identifiable information was removed from the transcripts ahead of analysis.
- To understand the capabilities of the LLM to conduct thematic analyses with minimal researcher involvement, we outlined the study objective before leading with a broad initial prompt. The basis of the prompt is presented below:
  - *"Imagine you have conducted interviews that seek to understand the patient journey, with a specific focus on the burden of disease for patients, their family friends and carers, alongside the burden of disease on the NHS. These interviews have been conducted to inform future research for a novel therapy. Please analyse these transcripts using thematic analysis to highlight the key insights emerging"*
- Following this, researchers guided the LLM iteratively through prompt development, to tailor the results to obtain more specific insights that the LLM may have initially missed. This included but was not limited to:
  - *"Highlight differences across key stakeholder transcripts e.g. those named GP or haematologist versus those named psychologist"*
- Each transcript also underwent manual qualitative analysis by an experienced researcher, who used validated and widely practised techniques used within thematic analysis and coding.
- LLM-generated themes were reviewed against the manual analysis for alignment and accuracy.
- To explore the potential scalability of AI analysis, the LLM was also tested on 20 transcripts.

Results

- With the initial prompt, the LLM accurately identified key themes such as epidemiology, clinical burden, and treatment pathways, with minimal revision.
- However, it struggled to interpret nuanced language conflating opposing mortality expectations and overlooking distinct psychosocial impacts.
- For instance, when describing the psychosocial/socioeconomic impact, theme generation was limited to describe domains it had identified which were particularly impacted. However, when contrasted with the manual analysis, several themes were overlooked, including the impact of the domain on a patient's ability to perform activities of daily living. Additionally, it was unable to provide wider context, which was identified by the researcher, including the identification of possible inaccuracies within the literature relating to the extent of burden experienced by patients (**Table 1**).

Table 1. Comparison of themes identified by manual analysis and AI

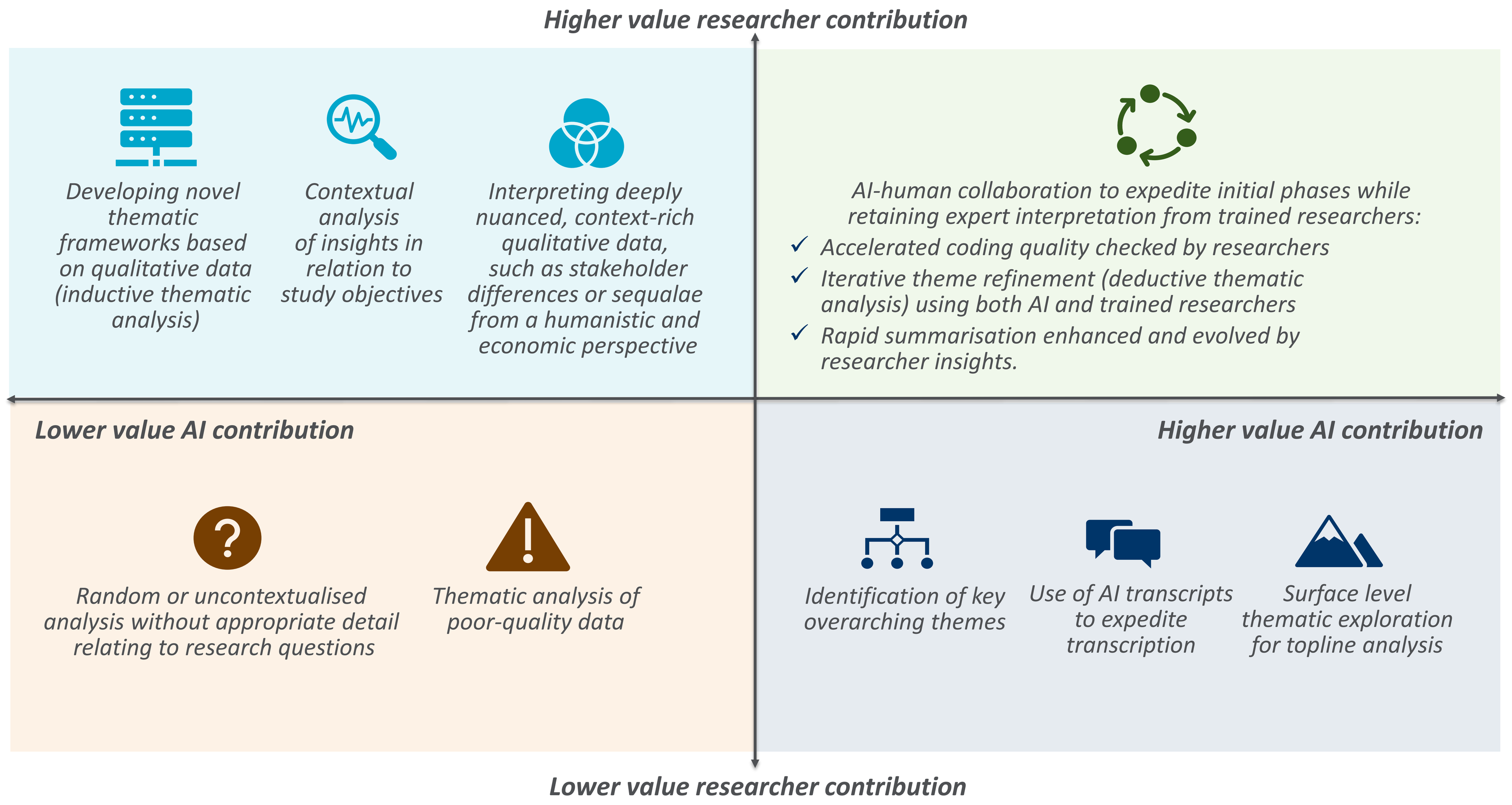
✓ Themes identified by AI	✗ Themes identified by researchers that were not identified by AI
Description of epidemiology, prevalence data and patient demographics.	Nuance relating to inaccuracies in prevalence data.
Symptom burden and quality of life with a focus on key symptoms and their burden.	Increased severity and urgency required during crises.
Description of frequency of complications and a focus on the cost of end organ damage.	The economic impact of complications when considering HCRU and the value of a novel therapy.
Diagnosis and screening with a focus on newborn screening and the risks associated with late diagnosis.	Challenges with diagnosis and screening including patient cohorts whereby existing prevalence data may not represent the true prevalence of the condition as a result as current diagnostic practices.
Treatment pathway and standard of care focusing on prevention and management.	Health disparities when considering the accessibility and availability of specialist centres.
Description of the transition from paediatric to adult care and the vulnerability of the period.	The importance and involvement of caregivers within paediatric and adolescent patients, alongside disparities relating to the intensity of monitoring from paediatric to adult care.
Psychosocial/ socioeconomic impact describing domains where burden is particularly high.	The impact on patients' ability to perform activities of daily living and potential inaccuracies in literature relating to the extent of burden experienced by patients
Unmet needs and barriers primarily focused on the lack of curative of highly effective treatment options.	The need for increased HCP awareness and education, expertise and training, and lack of multidisciplinary team involvement in disease management.
Economic burden with a focus on key drivers of direct costs associated with the condition.	Key drivers and factors conflating the cost associated with the management of complications, alongside the long-term impact of severe complications from an economic perspective.

AI: Artificial intelligence; HCP: healthcare professionals; HCRU: healthcare resource utilisation

Results (continued)

- While the outputs from the model provided a useful foundation, the outputs lacked significance when considering the application of thematic analysis findings in the context of the study objectives.
- Researcher expertise was required to synthesise outputs when considering key challenges for a novel therapy in the proposed indication.
- When scaled to 20 transcripts, the model necessitated segmentation to manage input length. This segmentation, however, introduced hallucinations and inconsistencies, raising concerns regarding the model's ability to synthesise large qualitative datasets.
- An illustrative overview of the key findings of this pilot relating to high value uses of AI and researcher contribution can be found below in **Figure 1**.

Figure 1. Comparison of themes identified by manual analysis and AI

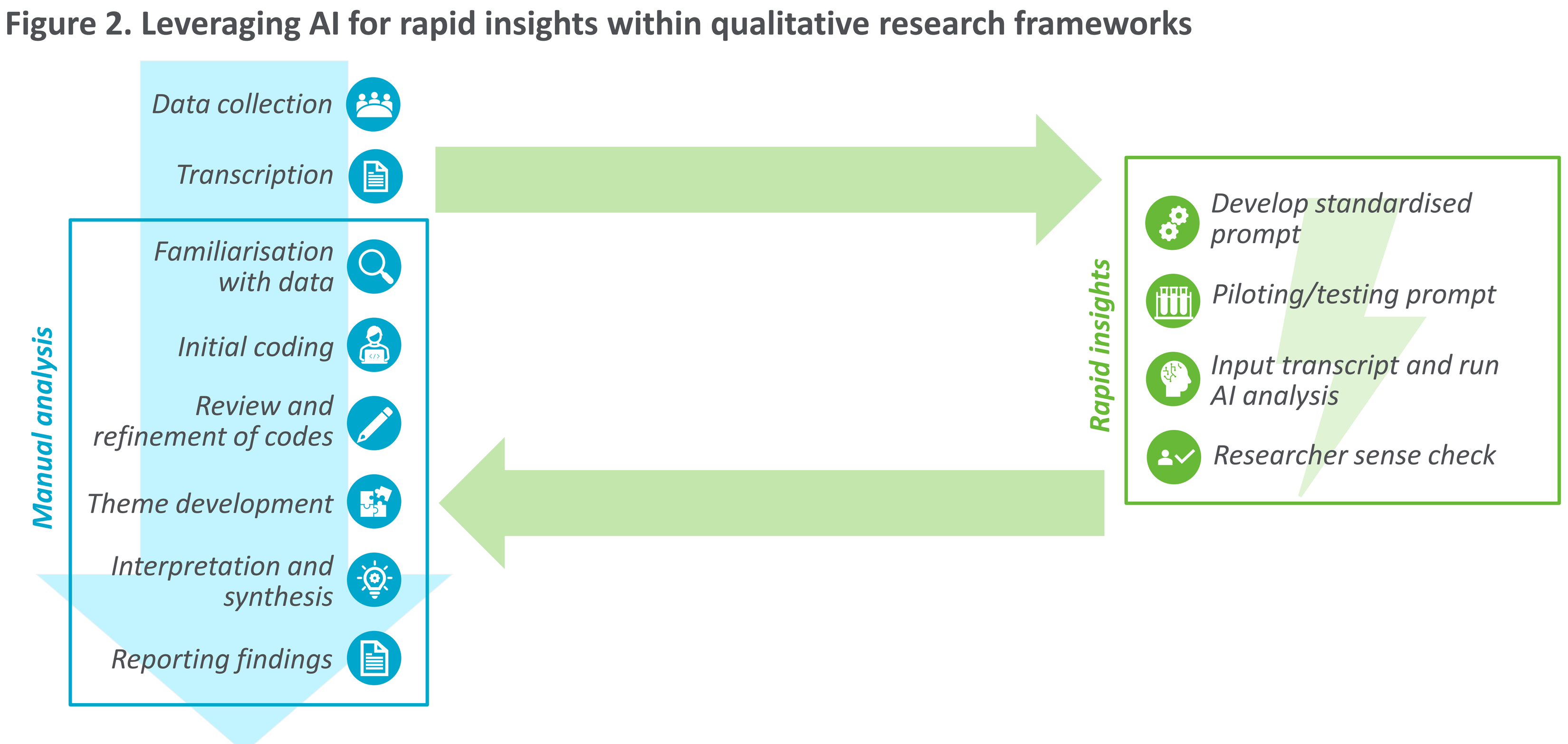


AI: Artificial intelligence

Integrating AI into a qualitative research process

- Our results suggest that relying solely on AI for qualitative research is not currently feasible; however, there is a significant opportunity to leverage AI for rapid insights.
- In **Figure 2** we provide an outline of the proposed process flow to allow for the integration of AI for rapid analysis, using a hybrid approach with validated manual analysis methodology.
- AI-driven analysis can efficiently process data sets, such as interview transcripts, to identify key themes.
- The use of standardised prompts and tailored queries enables consistent and repeatable extraction of insights across multiple transcripts or datasets.
- The insights generated can be reviewed to highlight top-line trends, supporting early-stage hypothesis and informing the need for alterations to questioning (e.g., during interview pilot phases).
- The outputs generated by AI can be integrated into the manual analysis framework to allow for cross-validation and enhanced rigor, with manual researchers ensuring that the themes they have generated aligns with AI outputs.
- Our approach has highlighted that AI may be useful to identify key themes preliminarily ahead of researcher involvement. When considering different types of thematic analysis, such as inductive (whereby the model identified themes and patterns directly from data) and deductive analysis (whereby the model applies predetermined themes from existing theories to analyse data), the value of AI may vary.
- A hybrid approach should be adopted across both instances. For inductive thematic analysis, it may be beneficial to leverage AI to identify key overarching themes that are them iteratively adapted, advanced and refined by researchers. For deductive thematic analysis, building the predetermined themes into a well-defined prompt may also serve as a useful option to help expedite the development of the coding process.
- In either scenario, the use of a hybrid approach accelerates the initial stages of qualitative analysis and allows for a greater level of collaboration between the research team and sponsor.
- By thoughtfully integrating AI-enabled rapid insights with established qualitative methodologies, there is an opportunity to deliver robust, timely, and actionable results that are tailored to the unique requirements of each project.

Figure 2. Leveraging AI for rapid insights within qualitative research frameworks



AI: Artificial intelligence

Conclusions

- LLMs can accelerate thematic analysis and facilitate the integration of key stakeholder insights into evidence submissions, potentially expediting evidence generation. However, in our exploratory use case, the LLM struggled to capture the nuances of the qualitative data, producing more basic outputs that lack depth, and struggles to synthesise large datasets without expert oversight.
- Human expertise remains essential for prompt refinement and contextual interpretation of qualitative stakeholder insights. As AI models evolve, integration within qualitative research may offer meaningful opportunities to streamline analysis. However, it remains important to implement the use of AI in such analyses with structured guidance and expert oversight. Future research should focus on developing scalable approaches to maintain accuracy across larger transcript sets.

References

1. Braun V, Clarke V. Thematic analysis. APA handbook of research methods in psychology, Vol 2: Research designs: Quantitative, Qualitative, Neuropsychological, And Biological., 2, 57–71. Washington, DC, US: American Psychological Association. 2012; **2.** Denzin NK, Lincoln YS. The Sage handbook of qualitative research. sage; 2011; **3.** Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qualitative research in sport, exercise and health. 2019;11(4): 589-597; **4.** Barrett D, Twycross A. Data collection in qualitative research. Royal College of Nursing; 2018. p. 63-64; **5.** Bazen A, Barg FK, Takeshita J. Research techniques made simple: an introduction to qualitative research. Journal of Investigative Dermatology. 2021;141(2):241-247. e1; **6.** Nowell LS, Norris JM, White DE, Moules NJ. Thematic analysis: Striving to meet the trustworthiness criteria. International journal of qualitative methods. 2017;16(1):1609406917733847; **7.** Pope C, Ziebland S, Mays N. Analysing qualitative data. Bmj. 2000;320(7227):114-116; **8.** Dwivedi YK, Hughes L, Ismagilova E, et al. Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. International journal of information management. 2021;57:101994.