The Perils and Pitfalls of Artificial Intelligence in Qualitative Interviews

Norzaris Bin Abdul[1], Balvinder Kaur Kler[2]*, Siao Fui Wong[3]

[1] Faculty of Business, Economics & Accountancy, Universiti Malaysia Sabah, Kota Kinabalu, Sabah, MALAYSIA.
[2]* Borneo Tourism Research Centre, Faculty of Business, Economics & Accountancy, Universiti Malaysia Sabah, Kota Kinabalu, Sabah, MALAYSIA.
[3] MODUL School of Tourism and Hospitality Management Nanjing, Nanjing Tech University Pujiang Institute, Nanjing, Jiangsu, CHINA.

Corresponding Author's email: balvinder@ums.edu.my

ABSTRACT

This paper critically evaluates the use of Artificial Intelligence (AI) tools for qualitative interviews conducted for an Interpretative Phenomenological Analysis (IPA) study. While AI-powered transcription and coding functions, such as those integrated into software, are efficient, for example, in ATLAS, this study highlights significant limitations in their ability to capture the emotional richness, contextual nuance, and reflexive insights that are central to IPA. Based on personal experience using both AI and, later, manually transcribing over 20 hours of semi-structured interviews with local champions in community-based tourism (CoBT), it became increasingly clear that overreliance on AI tools risks producing fragmented, superficial interpretations. The purpose of this paper is to discuss the experience of using AI to record, transcribe and analyse qualitative interviews. This paper presents comparative examples of AI-generated and human-interpreted code, illustrating that any meaningful understanding of participants' lived experiences requires more than technical efficiency. Manual transcription, which requires repeated listening and reflective note-taking, enables deeper immersion in the data, ultimately strengthening the interpretative depth and methodological integrity of the analysis. Findings reinforce the importance of the researcher's active and empathetic role in comprehending complex human narratives, a quality current AI tools are not equipped to replicate. As such, this study contributes to ongoing scholarly conversations surrounding the responsible use of AI in qualitative research. AI should be seen as a supplementary tool that enhances, rather than replaces, the essential critical and interpretative engagement inherent in qualitative inquiry, particularly within IPA. It is important to recognize that human insight and reflexivity remain central to this research approach.

Keywords: AI tools, Interpretative Phenomenological Analysis, qualitative interviews, transcription

INTRODUCTION

Artificial Intelligence (AI) is reshaping the landscape of qualitative research by expediting data collection and analysis. In research, AI and automated transcription tools, such as Otter.ai, accelerate research efficiency (Samuel & Wassenaar, 2025); however, integrating AI into the conduct and analysis of qualitative interviews raises significant methodological and ethical concerns. One issue is its limited ability to conduct in-depth evaluations (Aggarwal & Karwasra, 2025). Scholars warn that the use of AI in research could lead to illusions of understanding, in which researchers mistakenly believe they have fully comprehended the quick, convincing knowledge provided by AI (Messeri & Crockett, 2024).

These concerns are particularly relevant in qualitative interviews, which demand more than just accurate transcription. Qualitative interviews are data collection methods that use either open or semi-structured questionnaires to explore participants' subjective perspectives, meanings, and experiences (Hopf, 2004). Success largely relies on participants' spoken word, emotions, and visual cues, which are critical for increasing interpretive depth (Ahmad Azhari et al., 2022; Chong, 2022). Kowal and O'Connell (2004) argued that transcription is not simply just description or coding; interpretation depends on what and how narratives are expressed and incorporated into transcription.

Unlike quantitative surveys, qualitative interviews depend on context, empathy, and interpretive depth (Chong, 2022; Hopf, 2004), human qualities that AI cannot replace. The importance of being 'as human as possible' during interviews is to increase emotional connection and narrative quality (Ng et al., 2022). Qualitative

interviews require self-reflexivity (Chong, 2022), but using AI in transcription risks reducing complex narratives into decontextualised data. Also, the role of the researcher-as-instrument is essential to the entire IPA process, as the double hermeneutic circle empowers the researcher to interpret experiences (Smith & Osborn, 2003). Therefore, should a researcher's 'power' be replaced by AI?

This paper was motivated by the experience of applying the AI tool Transkriptor in a qualitative study. The purpose of this paper is to discuss the perils and pitfalls of using AI in qualitative interviews, arguing that AI's convenience compromises the interpretative depth of qualitative research. Deliberation begins with an overview of AI in qualitative research, followed by an evaluation of AI and its role in interviews. Next, the importance of transcription in an Interpretative Phenomenological Analysis (IPA) study is presented, followed by a delineation of the method. Subsequently, a section titled "the experience of AI in qualitative interview transcription" presents and juxtaposes a comparative analysis of AI-generated versus human-refined qualitative codes. The discussion section evaluates AI's experience in transcribing qualitative interviews and presents propositions before concluding the paper.

BACKGROUND OF STUDY
Artificial Intelligence in Qualitative Research
Artificial Intelligence (AI) is defined as the ability of autonomous systems to interpret their environment to achieve specific goals (Sheikh et al., 2023). AI has been used in qualitative research since the 1980s through tools such as Computer-Assisted Qualitative Data Analysis Software (CAQDAS), including NVivo and ATLAS.ti (Nguyen-Trung, 2025). In the past, these tools that integrated machine learning were limited by predefined tasks or algorithms and lacked meaningful interaction with human researchers (Nguyen-Trung, 2025).

Over the years, AI applications in qualitative research have evolved from early CAQDAS tools toward more sophisticated systems. In recent years, the rise of Generative AI (GenAI), such as ChatGPT and DeepSeek, which are integrated into qualitative data analysis tools, has been transforming qualitative research. These GenAI are used to assist transcription, summarise data, and produce code or themes. This evolution highlights the expansion of AI tools from supporting basic categorisation to complex pattern recognition and semantic analysis.

Presently, AI is widely applied in qualitative research in different stages. Deep learning models like ChatGPT can assist literature reviews by summarising large volumes of research and identifying key themes, which supports conceptual research aimed at generating ideas and identifying key concepts (Christou, 2023). In data collection and analysis stages, AI enables automated transcription of interviews and focus group data, conversion of audio or video files into text, thus saving time and reducing costs (Marshall & Naff, 2024; Samuel & Wassenaar, 2025)—AI-assisted coding and thematic analysis, which are made common in tools like ATLAS. Ti also shortened the time for data organisation and analysis (Nguyen-Trung, 2025).

While these developments reflect increasing AI integration in the research process, they represent a significant change from earlier use, which was primarily mechanical and descriptive. For instance, Marshall and Naff (2024) highlight AI's ability to do the following: (i) increase efficiency by speeding up the time-consuming transcription process; (ii) reduce cost through automated transcription and analysis; (iii) reduce manual labour cost; (iv). improve accessibility by supporting text conversion from audio or video; (v) support data analysis using data coding and analysis tools such as ATLAS.ti and (vi) generate ideas and answer complex questions using ChatGPT.

One study, which used Guided AI Thematic Analysis (GAITA) to examine the use of ChatGPT in qualitative data analysis (QDA), also found ChatGPT beneficial in increasing coding efficiency, data organisation, and theme generation (Nguyen-Trung, 2025). Likewise, Samuel and Wassenaar (2025) also acknowledged the advantages of AI in data transcription. It can be surmised that AI has transformed from a passive analytical tool to an active assistant for data interpretation and the research process. Spanning the process from literature review to data transcription, organisation, and analysis, AI enhances efficiency and supports multiple stages of qualitative research. Despite the benefits, this paper demonstrates that AI remains limited in both transcription accuracy and interpretive depth, particularly in non-English-speaking contexts.

In qualitative research, depth of interpretation is crucial. Indeed, Marshall and Naff (2024) caution that overreliance on AI risks losing human elements, leading to oversimplification of data interpretation. Undeniably, reliance on AI may streamline the research process but at the cost of illusions of understanding, which greatly reduces originality and critical reflection (Messeri & Crockett, 2024). Scholars have warned that the reliance on AI risks producing less capable researchers who lack critical thinking and reflexive ability (Marshall & Naff, 2024; Messeri & Crockett, 2024). Christou (2023) highlights that AI, particularly those that use deep learning

models, can produce biased responses because AI is trained on biased data. Without proper supervision, these AI-generated outputs may cause serious issues, while the complexity of deep learning models can make their outputs difficult to interpret (Christou, 2023).

Similarly, Aggarwal and Karwasra (2025) argue that while AI improves research efficiency, it often lacks empirical evidence, contextual understanding, and theoretical depth, and cannot replace human researchers' critical thinking and domain expertise. Although Nguyen-Trung (2025) also maintains that AI improves research efficiency, human intelligence remains essential for ensuring rigorous documentation of the research process and reflexivity, which are fundamental to qualitative research. Moreover, an IPA study is committed to the double hermeneutic circle, in which "the researcher is trying to make sense of the participant trying to make sense of their experience" (Smith & Osborn, 2003, p. 35). Can AI replace the researcher-as-instrument, and if so, to what extent?

Notably, these functional developments are accompanied by ethical implications of integrating AI into qualitative research. Generally, AI raises a broader ethical issue by fostering overconfidence among researchers. As Messeri and Crockett (2024) highlight, overreliance on AI can create illusions of understanding, thereby exploiting human cognitive biases. Without proper ethical guidelines, AI in research can be harmful. For instance, while Samuel and Wassenaar (2025) acknowledged the advantages of AI in data transcription, they also highlighted ethical concerns related to informed consent, data privacy and security, and the potential inaccuracy of AI-generated transcripts. Such ethical challenges were also highlighted by Nguyen-Trung (2025), who emphasised the need for a detailed record of ethics applications that address issues of data ownership and privacy. The on-going debate continues to question whether AI in qualitative research enhances efficiency or reduces interpretative depth and ethical integrity.

Role of Artificial Intelligence in Qualitative Interviews

While AI is widely used in qualitative research, its role in interviews is still emerging (Marshall & Naff, 2024). Marshall and Naff (2024) are optimistic about the potential future use of AI, such as ChatGPT or chatbots, for conducting and simulating interviews like human interviewers. However, concerns have been raised about ethical considerations and limitations in achieving the depth and contextual understanding that can only be derived through human interpretation. At present, AI has already established its presence in specific processes of qualitative interviews, such as transcription, coding, thematic analysis and interview facilitation and interaction. This section explores recent, specific advancements in AI applied to qualitative interviews, distinguishing them from general use.

First, during transcription, tools such as Otter.ai can greatly reduce time and cost by converting audio or video into text, either through uploading audio or using built-in recording or transcription functions in online videoconferencing platforms (Samuel & Wassenaar, 2025). Although these tools enhance efficiency, ethical concerns regarding informed consent, data privacy, and the potential risks of handling non-de-identified data (i.e., data that still contains personally identifiable information) have been highlighted (Marshall & Naff, 2024; Samuel & Wassenaar, 2025). Furthermore, Marshall and Naff (2024) warned that AI transcription may limit researchers' early engagement with the data, a crucial step for developing familiarity and gaining an in-depth understanding of raw data in qualitative analysis.

Secondly, in coding and thematic analysis of interview data, AI tools such as ChatGPT offer quick searches of predefined codes, allowing (i) extraction of relevant quotes; (ii) supports pattern identification and narrative construction; and (iii). has helped researchers perform deductive thematic analysis, significantly reducing labour and time required in manual coding (Chubb, 2023). Similarly, AI features built into data analysis software like ATLAS.ti facilitate the coding process but raise concerns, including treating AI-generated codes or themes as final results, which contradicts the flexible, interpretive nature of qualitative research (Marshall & Naff, 2024). Additionally, GenAI tools like ChatGPT further enhance qualitative interview analysis by accelerating initial coding, visualising data, and generating preliminary themes. However, they cannot replace the researcher's interpretative and critical thinking abilities, which are critical for refining results and maintaining methodological rigour (Nguyen-Trung, 2025). Such advancements demonstrate the growing reliance of AI in the interview process, even as researchers remain central to meaning-making and ethical interpretation.

As mentioned earlier, AI tools have not been widely used to conduct qualitative interviews. This concern stems from the risk that AI may reduce active participation and the relational nature of qualitative interviewing, potentially diminishing data richness (Marshall & Naff, 2024). This issue is particularly important as the success of qualitative interviews relies heavily on the 'what and how' participants speak of the phenomenon (Ahmad Azhari et al., 2022; Chong, 2022; Kowal & O'Connell, 2004). However, current research has acknowledged AI's specific role in facilitating interviews and interactions. For example, Nardon et al. (2025) found that AI image

generation supports participants and interview interaction as a helper, motivator, and facilitator when it is not a distractor and influencer. By considering an AI image generation tool as a third agent in the interview process, Nardon et al. (2025) praised AI's ability to enrich the interview process by adding "new layers of content and meanings to the dialogue" (p. 9). Additionally, the ChatPDF tool facilitates researcher interaction with qualitative data through prompts or prompt engineering (the input or instructions given to the AI system), but its effectiveness depends on the researcher's ability to provide instructions (Chubb, 2023).

Although AI tools are beginning to transform qualitative interviews, their use comes with significant ethical and methodological concerns. Unlike human interviewers, AI lacks the capacity for spontaneous, empathetic interaction and contextual probing, which are critical for gathering rich data (Marshall & Naff, 2024). Transcription, in particular, requires the inclusion of participants' emotions to comprehend the meanings behind narratives fully.

This underscores the need for researchers to remain reflexive and critically engaged to enhance interpretative depth and uphold ethical integrity (Nardon et al., 2025). Addressing unresolved methodological and ethical concerns, Agarwal and Karwasra (2025) advocate expanding AI applications beyond current use cases while seeking a balance between AI and human intelligence in academic research. Similarly, Nguyen-Trung (2025) stresses the importance of reflexivity and cautions against allowing AI to supplant the interpretative depth essential to qualitative research, warning of risks such as data bias and superficial analysis. Likewise, Nardon et al. (2025) call for ongoing dialogue to ensure that AI is merely a supportive tool in qualitative research. Together, these scholars highlight a critical research gap, the need for careful investigation into the perils and pitfalls of integrating AI into qualitative interviews in ways that safeguard the interpretative nature of qualitative research. The following section focuses this discussion on the use of AI within a specific qualitative approach.

IPA Interviews and the Importance of Transcription

Interpretative Phenomenological Analysis has theoretical underpinnings in phenomenology, hermeneutics and idiography (Smith et. al, 2022). Seeking to understand lived experiences, IPA involves a detailed examination of each participant's narrative to understand how they make sense of their personal experiences. This idiographic focus requires rich, detailed accounts from a small purposive sample, typically 5 to 10 participants. In IPA, each 'case' is analysed thoroughly before a cross-case comparison is conducted to identify themes that are common across participants. IPA encourages the emergence of unexpected themes from the data, emphasising participants' role as experts in producing thick, rich descriptions of their lived experiences, which are essential for phenomenological interpretation.

In IPA, semi-structured interviews are the primary method for assessing participants' lived experiences, as they enable participants to express their experiences in their own words with minimal directive questioning (Smith et al., 2021). Hence, physical, temporal, and interpersonal dynamics of interviews play a key role in shaping the depth and quality of the data gathered. A fundamental aspect of IPA's rigour is transcription, which goes beyond simply converting narratives into text. As is common in other qualitative approaches, the interpretative process requires capturing pauses, tone, and emphasis to preserve the authenticity and emotional context of participants' narratives (Kowal & O'Connell, 2004). Here, the ability to listen and capture emotional cues during transcription not only produces high-quality transcription but also supports reflexive, deeper engagement with participants' meaning-making processes.

It needs to be noted that interviews rely on context, empathy, and researcher reflexivity, human qualities that embrace the 'art of listening' (Bong, 2002), are not imitable by AI (Chong, 2022; Hopf, 2004) and crucial for trustworthiness of the findings (Ng et al., 2022). Although AI can assist with transcription and coding, this paper proposes that it cannot capture the subtle interpretative nuances required for IPA. Also, overreliance on AI tools risks overlooking the depth and context needed to understand participants' lived experiences truly. For this reason, this paper emphasises the continued need for human-led transcription and reflective note-taking to preserve the interpretative depth and rigour demanded by IPA. To illustrate these challenges, the following section delineates the research context, data collection and transcription, drawing on experience with one AI transcription tool, Transkriptor.

METHOD

This section outlines the steps taken to collect and transcribe data for an IPA study, contextualising the challenges arising from the use of AI, specifically drawing on experience with Transkriptor, a built-in transcription feature in ATLAS.ti (version 24).

Research Context

In a CoBT study, the first author conducted face-to-face, semi-structured interviews to explore the lived experiences of twelve local champions recruited from various locations in the West Coast of Sabah, Malaysian Borneo region (see Abdul et al., 2024). CoBT is a popular model in rural tourism that supplements the income of rural agricultural communities, in which cooperatives own and manage tourism enterprises (Lo & Janta, 2020). An important factor in the sustainability of CoBT is the role of a 'local champion' who keeps the community motivated by focusing on long-term goals that include preserving culture, heritage, and the environment (Ginanjar et al., 2024). The primary research question asked "how do local champions experience CoBT" with two objectives: i. to describe the lived experience of CoBT for local champions, and ii. to illustrate the meaning of CoBT for local champions. This context is used as a backdrop to the discovery of what AI lacked during data collection and analysis in the CoBT study.

Data Collection

Interviews were conducted between September 2023 and June 2024 in various locations (e.g., traditional community houses, homestay porches, office spaces, cafés) chosen by the participants and convenient to them. These settings were used to ensure the researcher could record and observe participants' speech, including emotional expressions, pauses, speed, tone, and cultural references. A recording device was imperative during data collection. Even though the first author was there in person, it was challenging to capture all these subtle cues accurately. Upon completion of data collection, the first author personally transcribed each interview carefully to prepare for IPA analysis. During transcription, it was imperative to capture all subtle cues, pauses, emotional expressions, and non-verbal cues noted during the interview. In the transcription process, an important step in IPA is to listen to the recordings repetitively. This step ensures that a researcher becomes familiar with the data and transcribes it accurately. The following section provides a detailed account of the process.

Transcription and Analysis

The initial phase of data processing was accelerated through the use of AI transcription tools, namely Transkriptor, a feature integrated into the ATLAS.ti software. This tool was chosen for its seamless compatibility with the planned data analysis workflow. The process began with preparing audio files, during which all interview recordings were uploaded to ATLAS.ti in MP3 format to ensure consistent audio quality. A transcriber was then used to automatically convert speech to text, typically producing initial transcripts within 15 to 30 minutes per hour of audio. After that, the first author conducted an initial review of the AI-generated transcripts, identifying and noting any apparent errors or inaccuracies for subsequent correction. Next, the AI-based coding feature in ATLAS.ti was used to generate preliminary codes and themes from the transcribed data.

During transcription, the researcher had to listen carefully, repeatedly, and immerse themselves in the recordings against any notes taken on-site. This situation would have been difficult for AI tools to capture as A.I often struggle to fully and precisely understand the richness of human speech. More so when local accents are involved and the language spoken is colloquial. Indeed, several significant challenges arose during this process. Transkriptor struggled with multilingual content, especially when participants switched between English, *Bahasa Melayu,* and local languages, leading to unclear or incomplete transcripts that required significant manual correction. The tool also failed to accurately transcribe cultural terms, local place names, and traditional concepts that were important to the study. These had to be added manually or revised in text. In addition, the AI was unable to capture essential paralinguistic elements such as pauses, changes in tone, emotional inflections, laughter, or silence, all of which are crucial in an IPA study. Technical limitations, such as background noise in natural interview settings, also affected transcription accuracy, as the tool struggled to distinguish between speakers' voices and environmental sounds. Therefore, the first author spent considerable time revising the transcripts.

Upon completion of transcription, all transcripts were once again carefully reviewed multiple times. This step ensured familiarity with the data. In the words of Smith et al. (2022), this repetitive examination of transcripts is a process of "dwelling with the data", where the researcher immerses themselves in the participant's stories by reading the transcripts while listening to the original recordings at the same time. This approach ensures the researcher listens attentively and captures subtle meanings, emotions, and nonverbal cues that may be overlooked when reading the transcript alone. Smith et al. (2022) call this deep immersion "slowing down our habitual propensity for 'quick and dirty' reduction and synopsis" (p. 82). In this phase, while reading and listening, the researcher also recorded initial impressions. At the same time, attempts were made to identify recurring themes from the data and mark (or code) unique expressions in the transcript.

This process indicates a transition from data collection to data organisation and examination. The marked transcripts then served as the basis for a thorough and individual analysis of each interview's data. This process

aligns with a key characteristic of IPA, which requires a close examination of each participant's experiences before moving to the next. As such, this process ensured that the researcher fully understood each participant's unique story and perspective. By doing so, the richness and detail of each participant's lived experience were preserved. We propose that none of these steps is achievable with AI tools, as elaborated in the next section under 'Findings.'

FINDINGS
The Experience of AI Transcription
This section outlines the experience of using AI to generate qualitative codes and explains why the decision was made to regenerate the codes manually. Initially, the first author attempted to use an AI transcription tool to expedite the conversion of the over 20 hours of interview recordings into text. The researcher employed Transkriptor to perform the transcription. Indeed, the tool was efficient at converting recordings to text. However, soon several significant perils and pitfalls of the tool became apparent, significantly reducing its effectiveness for IPA inquiry.

First, it struggled with transcribing recordings that were not in English. For the study, interviews were conducted among local people from Sabah, Malaysian Borneo. Participants were given the option of being interviewed in English or *Bahasa Melayu*. Notably, although a participant chose to speak *Bahasa Melayu* in Sabah, the national language is specifically referred to as a 'Sabah Malay Dialect' (SMD) (Wong, 2012). Often, participants mix two languages when speaking and even use words from native dialects in the same sentence. Therefore, Transkriptor frequently produced incoherent and inaccurate transcripts. It was time-consuming to correct these errors over conducting manual transcriptions.

Secondly, AI tools were unable to capture emotional and nonverbal cues when SMD was spoken. Even with English narratives, AI tools still failed to adequately capture pauses, tone, emphasis, emotional expressions, and nonverbal sounds, which are fundamental to phenomenological research. These significant limitations highlight a clear pitfall of AI tools and reveal a deeper peril: reliance on AI for transcription significantly reduces researchers' reflexivity and engagement with participants' experiences. This creates a gap between the researcher and the data, which contradicts the notion that transcription should involve deep researcher engagement with participants' experiences (Smith et al., 2022). Thus, AI causes more than technical problems. It poses a danger to the integrity of IPA research.

At the peer review stage, the supervisory critique indicated that the data were dry and emotionless, lacking meaningful descriptions of lived experiences expected in an IPA study. Themes produced lacked the flavour of experiential CoBT moments that would highlight the value of being a local champion. The supervisory team advised that the analysis be revised manually without the use of AI. Therefore, the experience of using AI for transcription led the first author to realise that AI did not increase immersion in the data. Instead, it turned the participants' experiences into mechanical, superficial texts, reducing them to decontextualised data. To show the differences between human-interpreted and AI-generated codes, Table 1 lists three sample interview excerpts, with A. I assigned codes and human interpreted codes and a final column with critique on the AI coding. Here, a deeper peril is revealed: the detachment from participants' lived experiences and emotions.

As shown in Table 1 (see below), Participant 1 clearly struggled to understand how his personality contradicts the fact that he was chosen by the community to lead, stating, "*yet, puzzlingly, they chose me to lead them. I couldn't comprehend why.*" This was an emotional and thematic centre of the narrative, indicating the participant's confusion between how he saw himself and how others saw him. The underlying message reads "Why would people choose someone like me—different, unsociable—to lead?" This narrative, coded by AI as conflict, isolation, leadership, self-reflection, and social withdrawal, was accurate but consisted of fragmented descriptions that identified individual traits or themes. These codes failed to connect the separate elements into a meaningful synthesis to interpret the participant's confusion and capture emotional moments of self-reflection and deeper self-understanding. The first example shows that AI coded risks as being shallow.

In the second example, Participant 8 shared a story about taking a substantial financial risk, and afterward someone said, "He said it was good enough." This short final sentence means the world to the participant because it indicates both approval and evidence that the risk paid off. Codes generated by AI, such as entrepreneurship, financial struggles, growth mindset, resilience, and risk-taking, described a general situation but missed the emotional climax —the moment when the participant felt accepted, and finally, his effort and sacrifice were valued and acknowledged. This sentence, which shows the participant's journey from 'taking a risk' into a story of 'success, recognition, and personal achievement', represents a powerful moment that captures the true meaning and

emotional impact of the participant's experience, but was missed by AI. This example shows an incomplete, flat analysis of AI coding. Here, a perilous consequence of automated coding is the risk of failing to capture genuine emotion.

Table 1: A Sample Comparative Analysis of AI-Generated vs. Human-Refined Qualitative Codes

Participant	Interview excerpt	AI-Assigned Codes	Human-interpreted Codes	Critique on AI Coding
P1	*Admittedly, I have a fiery temper. People in this village know very well that I am not one to easily socialize. Typically, I play football and then retreat back to my home. I'm different from the rest, rarely mingling, **yet, puzzlingly, they chose me to lead them. I couldn't comprehend why.** When I was thrust into this significant role, leading the village community, it prompted me to reflect upon myself more.*	Conflict, Isolation, Leadership, Self-reflection, Social withdrawal.	Paradox of the Reluctant Leader; Leadership based on Action, Not Popularity.	Misses the Core Puzzle. The codes identify traits but miss the central paradox that puzzles the speaker.
P8	*I cleared all my bank accounts. I did it; he came and inspected. **He said it was good enough...***	Entrepreneurship, Financial struggles, Growth mindset, Resilience, Risk-taking.	All-In Risk, Validation after Sacrifice; Betting on Oneself.	Misses the Climax. The codes describe the process but miss the crucial moment of validation after the immense risk.
P11	*My husband and I developed this, and suddenly he fell sick. He had cancer; he is a strong man, a Hero! From a very healthy man to becoming very skinny and eventually passing away. So, for me, that was the most bitter experience I have ever had in this world. Ha...! That was the most painful. So maybe when I have a small problem, I don't feel sad anymore or worried because I have already faced the worst. That's my answer for that. I have been given the highest challenges that people can face. I have it; I faced it. If there's a small thing, it's nothing anymore to me. So yeah, that was it. That's an experience.*	Grief, Life challenges, Perspective shift, Resilience.	Transformative Loss; Grief as an Anchor; Redefined Hardship.	Clinically Detached. The codes are accurate but lack the emotional gravity of the experience described.

Similarly, an accurate but emotionally flat coding that significantly reduced the richness and impact of the findings was found in the third example. Participant 11 shared a personal and painful story of grief, loss, and transformation. While the AI coding accurately identifies key themes such as grief, life challenges, perspective shift, and resilience, it failed to capture the full emotional depth —the actual feelings of pain, sadness, and vulnerability of the participant.

This participant's story about watching a loved one get worse and eventually losing a loved one is not merely expressing tough times but is indicative of a deep emotional change that influenced her lifeworld: no other challenge would be insurmountable based on this experience of loss. AI coding used general labels and failed to express and explain the intensity of the personal experience.

This clearly shows that AI failed to reflect the complex mix of a human being's suffering and strength. Together, these examples demonstrate not only the functional pitfalls of AI tools but also the perils they pose to the interpretive integrity required in qualitative research.

DISCUSSION

As a doctoral student focused on completing transcription and analysis, after conducting these steps for all twelve participants, the emptiness and off-the-mark AI coding became obvious. This paper proposes that AI tools alone cannot adequately capture the richness of lived experience. Realising the issue, the first author began to acknowledge the importance of immersive engagement as a methodological necessity rather than a burden. For this reason, the researcher then committed to a complete manual transcription process. As mentioned, manual transcription involves multiple careful listenings to each recording to ensure both spoken words and non-verbal cues are captured. The analytical process followed Smith et al. (2022) and a tailored notion system was developed that included markers for pauses (duration for more extended silences were marked), emphasis (text that sounded like emphasis was underlined), emotional expressions and non-verbal sounds such as sighs or laughter (usually written in square brackets).

For interviews conducted in SMD, the researcher first transcribed the recordings in their original language to maintain linguistic authenticity. Then, transcripts were carefully translated into English for analysis. During the translation process, special attention was given to terms and expressions deeply rooted in local culture that cannot be translated directly into English. To avoid inaccurate translation, detailed translation notes were kept, and native speakers were consulted, including during the back-to-back translation process. This ensured an accurate capture of participants' meanings. During data analysis, both the original and translated transcripts were kept side by side to enhance accuracy and transparency. While this allowed for careful comparison of data between the two languages, it also enhanced accuracy and consistency.

Most researchers agree that manual transcription is time-consuming, and this is undeniable. However, such time investment has been proven invaluable for developing a deep familiarity with participants' experience. As Weitzman (2003) stated, qualitative researchers can "dump" the data into an AI tool and see what comes out. Table 1 shows what was produced and indicates that AI tools are incapable of capturing what is revealed through the slow, immersive engagement of a human researcher: the subtle patterns and emotional nuances that can only be understood by a human. For instance, the trembling of participants' voices when describing community conflict, the accelerated tone or pace when sharing their success stories, and the pauses before sharing a deep reflection; these are the obvious yet evident insights central to fully understanding participants' lived experiences. From the experience of AI-assisted and manual transcription, a notable difference emerges, highlighting the need for qualitative researchers to remain deeply engaged with the data —not just as a philosophical preference but as a methodological imperative. This paper proposes that relying on automated transcription leads to overlooking the very experiential aspects that phenomenology seeks to uncover.

Ultimately, the transcription process became instrumental in enabling researchers to fully capture the depth and richness of participants' biographies, situating each lived experience within a wider life context. Again, repetitive listening to recordings is necessary. Through this so-called 'time-consuming' process, time and effort are significantly reduced as researchers can identify recurring themes and references. In this study, for example, this element was evident in a participant's frequent references to her grandmother's words, which emerged as a key interpretative thread. As IPA focuses on idiographic commitment and the exploration of patterns across participants, these biographies balanced its focus by preserving confidentiality through pseudonyms and omission of identifying details.

This process highlighted a key epistemological difference between IPA and more distanced qualitative methods: whereas AI tools can assist thematic analysis with pattern detection, it lacks IPA's focus on the double hermeneutic circle, which requires a researcher's embodied engagement with data, to hear the voices of participants, and to feel participants' emotional resonance of their words (Smith et al., 2022, p. 84). Through this experience, the first author realised the move from relying on AI back to manual transcription had demonstrated that such closeness to data cannot be replaced without losing the deep phenomenological understanding IPA aims to achieve.

The experience of using AI transcription tools in this IPA study of CoBT local champions revealed significant impacts on data quality and theme discovery. In the initial phase, the principal author used Transkriptor, an AI tool integrated within ATLAS.ti, to accelerate data processing. However, several critical limitations quickly became apparent. The tool struggled with language and dialect barriers, especially with *Bahasa Melayu* and the SMD, and with frequent code-switching between English and local languages. This resulted in transcripts that were often unclear and inaccurate, sometimes requiring more time to correct than manual transcription would have taken. More importantly, AI failed to capture the emotional depth and contextual nuances essential to phenomenological research, such as pauses, tone, and emphasis, which are crucial for understanding participants' lived experiences.

Supervisory peer review highlighted that the themes generated by AI were "dry and emotionless," lacking the experiential richness expected in IPA. For instance, AI coding missed the emotional paradox in reluctant leadership, the pivotal moment of validation after risk-taking, and the complex interplay of suffering and strength in stories of personal loss. In contrast, manual transcription preserved linguistic authenticity and emotional context, allowing five cultural dimensions to emerge from the data: (i) *Misompuru* – the experience of fostering unity, (ii) *Pangazou* – the experience of weathering challenges, (iii) *Sumazau* – the experience of harmonizing traditions, (iv) *Sunduvan* – the experience of embodying stewardship, and (v) *Mitatabang* – the experience of building hope, which exemplify the lived experience of a local champion in CoBT.

The manual approach, described by Smith et al. (2022) as "dwelling with data," involved repeated listening, preservation of cultural context, and capturing emotional resonance, elements that AI tools could not replicate. Ultimately, this paper proposes that while AI transcription offers efficiency, it fundamentally undermines the interpretative depth and cultural nuance required for high-quality phenomenological research. A deep understanding of local champions' experiences, which is vital for sustainable CoBT, can only be achieved through immersive human engagement with the data.

LIMITATION OF STUDY

This paper was based on the experience with only one AI tool, Transkriptor, in version 24 of ATLAS.ti. While the combined use of AI and human interpretation provided evidence highlighting the perils and pitfalls of AI in our study, it is possible that the next version of the AI tool in ATLAS.ti or indeed other transcription tools, such as NVivo, might offer enhanced capabilities better suited to capture the meanings essential for IPA studies.

Future work could compare the effectiveness of various AI tools in transcription and analysis. The visible conundrum, suffice to say, is that applying AI yielded themes that lacked the depth needed for an IPA study. Illustrating the meaning of lived experiences for local champions was achieved through the epistemological depth provided within a qualitative approach. The peer review process was imperative for ensuring analysis and interpretation remained cognisant of immersion in the data and the double hermeneutic circle required in IPA. AI is convenient and hastens transcription and analysis, but lacks the closeness to data that the researcher-as-instrument must uphold.

CONCLUSION AND RECOMMENDATIONS

This paper highlights the critical limitations of using AI tools in qualitative interview data transcription and coding, particularly in conducting IPA. The paper acknowledges the role of AI tools, such as Transkriptor, in accelerating the conversion of recordings to text as well as the speed of ATLAS.ti in data coding. However, this paper critiques AI's limited capacity to capture subtle emotions and the deeper context needed to fully understand participants' lived experiences.

Through several examples, the experience presented above demonstrated that AI-generated code produces only fragmented descriptions with superficial meanings. The inability to capture emotional richness leads to the loss of deeper meanings. Instead, the comparison between AI and manual transcription and coding confirms that human-interpreted analysis can better grasp participants' lived experiences, thereby significantly improving the emotional richness and impact of the findings. This confirms that human involvement is crucial to ensure the trustworthiness and quality of qualitative data analysis.

Again, manual transcription, the integral 'time-consuming' process that involves careful listening to recordings multiple times before typing out each word and notable emotions, is an invaluable process for qualitative researchers. Not only does this process help develop familiarity with participants' experience, but it also allows researchers to meet IPA's requirements for in-depth interpretation, self-reflection, and critical analysis. The power

of the researcher-as-instrument interpreting experiences in a double hermeneutic circle is retained. Thus, this paper emphasises the importance of close, human-led transcription, coding, and reflective note-taking. Additionally, the use of AI tools raises ethical concerns like data privacy, informed consent, and the potential for overreliance. This highlights the irreplaceable and need for human oversight in maintaining the depth, trustworthiness, and ethical integrity of qualitative research.

Based on the propositions of this study, future work exploring the use of AI tools in qualitative inquiry should adopt a multidimensional approach that encompasses technological, methodological, and cultural aspects. Researchers are encouraged to conduct comparative studies across various AI transcription platforms (such as Otter.ai, Rev, and Trint) to evaluate their performance with multilingual data, especially in contexts involving code-switching and dialectal variation, while also establishing standardized accuracy benchmarks for phenomenological research. Sampling considerations should include diverse linguistic backgrounds, cultural contexts, and patterns of emotional expression to ensure that AI tool evaluation truly reflects real-world research conditions, rather than just controlled environments.

Language factors require particular attention, with future studies needing to examine AI performance across different language families, tonal languages, and culturally embedded expressions that carry phenomenological meaning beyond literal translation. Methodologically, researchers should develop hybrid approaches that combine the efficiency of AI for initial processing with the interpretative rigor of human analysis through systematic manual verification protocols.

Ethical frameworks must also be established for the use of AI in sensitive qualitative contexts, including issues of informed consent for automated processing, data privacy in cloud-based transcription services, and the potential for AI bias in cultural interpretation. Training programs should be developed to help qualitative researchers critically evaluate AI outputs, recognize when human intervention is necessary, and maintain the reflexive engagement essential to phenomenological inquiry. Finally, collaboration between qualitative researchers, AI developers, and cultural linguists is crucial for developing AI tools that are more responsive to the needs of interpretative research, while respecting the human elements that underpin phenomenological understanding and the researcher-as-instrument paradigm central to IPA methodology.

IMPLICATIONS OF THE STUDY

This paper contributes to the ongoing debate pertaining to the use of AI in qualitative research. This study echoes a common viewpoint shared by numerous scholars, that AI should be used as a supplementary tool rather than a replacement for the researcher's critical, empathetic, and interpretative role (Aggarwal & Karwasra, 2025; Marshall & Naff, 2024; Messeri & Crockett, 2024; Nardon et al., 2025; Nguyen-Trung, 2025). By doing so, qualitative research can embrace AI while maintaining its focus on a deep and meaningful understanding of human experiences.

Methodologically, this study highlights a need for clearer frameworks and ethical guidelines to oversee the integration of AI into qualitative inquiry. This includes ensuring transparency in reporting AI's role, safeguarding against algorithmic bias, and preserving researcher reflexivity to uphold rigour and credibility. Next, training has to be incorporated as AI becomes more embedded in academic practice. Future scholars will require both digital literacy and critical awareness to engage effectively with AI outputs without compromising interpretive depth.

Finally, this study carries practical implications for the wider academic community. Journals, institutions, and funding bodies will need to establish policies for responsible AI use, including disclosure requirements and ethical standards. At the same time, future work could explore the comparative value of AI-assisted versus human-only analysis to evaluate the validity, depth, and richness of insights. This could determine contexts in which AI enhances, rather than undermines, qualitative inquiry.

AUTHOR CONTRIBUTION

NA collected and analysed the data as part of a doctoral study. BKK conceptualised the idea for this paper. BKK and SFW verified findings via peer review and supervised the study. All authors composed sections of and refined drafts of the manuscript for publication.

DECLARATION OF STATEMENT
We declare that this is original research, and this article has not been published elsewhere.

ACKNOWLEDGMENTS

CONFLICT OF INTEREST STATEMENT
The authors declare that there is no conflict of interest regarding the publication of this paper.

REFERENCES
Aggarwal, V., & Karwasra, N. (2025). Artificial intelligence v/s human intelligence: a relationship between digitalization and international trade. *Future Business Journal*, *11*(1), 1–16. https://doi.org/10.1186/s43093-025-00438-5

Ahmad Azhari, A. E., Chai, J., & Anderson, C. (2022). Reflections of conducting online interviews during the pandemic: Benefits, challenges and practical solutions. *Malaysian Journal of Qualitative Research*, *8*(1), 24–31. https://doi.org/10.61211/mjqr080103

Bong, S. A. (2002, May). Debunking myths in qualitative data analysis. In *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research.* 3(2), Art. 10.

Chong, S. L. (2022). Interviewing in qualitative research. *Malaysian Journal of Qualitative Research*, *8*(1), 110–116. https://doi.org/10.61211/mjqr080110

Christou, P. A. (2023). How to use artificial intelligence (AI) as a resource, methodological and analysis tool in qualitative research? *The Qualitative Report*, *28*(7), 1968-1980. https://doi.org/10.46743/2160-3715/2023.6406

Chubb, L. A. (2023). Me and the machines: Possibilities and pitfalls of using artificial intelligence for qualitative data analysis. *International Journal of Qualitative Methods*, *22*, 1-16. https://doi.org/10.1177/16094069231193593

Ginanjar, R., Riani, A. L., & Aini, I. N. Q. (2024). Local champion: A new leadership perspective to create sustainable tourism in rural communities. The Eastasouth Management and Business, 2(02), 105-117. https://doi.org/10.58812/esmb.v2i02.196

Hopf, C. (2004). Qualitative interviews: An overview. In U. Flick, E. von Kardorff, & I. Steinke (Eds.), *A companion to qualitative research* (pp. 203-208). SAGE Publications.

Kowal, S., O'Connell, D. C. (2004). The transcription of conversations. In U. Flick, E. von Kardorff, & I. Steinke (Eds.), *A companion to qualitative research* (pp. 248-252). SAGE Publications.

Lo, Y.-C., & Janta, P. (2020). Resident's perspective on developing community-based tourism – A qualitative study of Muen Ngoen Kong Community, Chiang Mai, Thailand. Frontiers in Psychology, 11. https://doi.org/10.3389/fpsyg.2020.01493

Marshall, D. T., & Naff, D. B. (2024). The ethics of using artificial intelligence in qualitative research. *Journal of Empirical Research on Human Research Ethics*, *19*(3), 92-102. https://doi.org/10.1177/15562646241262659

Messeri, L., & Crockett, M. J. (2024). Artificial intelligence and illusions of understanding in scientific research. *Nature*, *627*, 49-58. https://doi.org/10.1038/s41586-024-07146-0

Nardon, L., Brüning, C., Valgardsson, S., & Busato, M. (2025). AI-image generation in research interviews: opportunities and challenges. *International Journal of Qualitative Methods*, *24*, 1-11. https://doi.org/10.1177/16094069251333335

Ng, Y. P., Kler, B. K., & Dousin, O. (2022). A reflexive account of fieldwork experiences: Interviewing tourist guides about the meaning of work during a pandemic. *Malaysian Journal of Qualitative Research*, *8*(1), 45-55. https://doi.org/10.61211/mjqr080105

Nguyen-Trung, K. (2025). ChatGPT in thematic analysis: Can AI become a research assistant in qualitative research? *Quality & Quantity*. https://doi.org/10.1007/s11135-025-02165-z

Samuel, G., & Wassenaar, D. (2024). Joint editorial: Informed consent and AI transcription of qualitative data. *Journal of Empirical Research on Human Research Ethics*, *20*(1-2), 3-5. https://doi.org/10.1177/15562646241296712

Sheikh, H., Prins, C., & Schrijvers, E. (2023). Artificial intelligence: Definition and background. In H. Sheikh, C. Prins, & E. Schrijvers (Eds.), *Mission AI: The new system technology* (pp. 15-41). Springer International Publishing.

Smith, J. A., & Osborn, M. (2003). Interpretative Phenomenological Analysis. In J. A. Smith (Ed.), *Qualitative psychology: A practical guide to methods* (pp. 53–80). Sage.

Smith, J. A., Flowers, P., & Larkin, M. (2022). *Interpretative phenomenological analysis: Theory, method and research* (2nd ed.). Sage Publications.

Weitzman, E. A. (2003). Software and qualitative research. In N. K. Denzin, & Y. S. Lincoln (Eds.), *Collecting and interpreting qualitative materials* (pp. 310–339). Sage.

Wong, J. K. L. (2012). *Sabah Malay dialect: Phonological differentiation in social context*. Unpublished doctoral dissertation. Jabatan Bahasa-bahasa Malaysia dan Linguistik Terapan, Fakulti Bahasa dan Linguistik, Universiti Malaya.