

# A comparative assessment of AI and manual transcription quality in health data: insights from field observations

SM Akramul Kabir, Fareeha Ali, Ruqayya Sulaiman-Hill

## ABSTRACT

**AIM:** This study explores the semantic similarities between qualitative research transcripts produced by artificial intelligence (AI) and those transcribed manually, with a particular focus on challenges encountered when working with multicultural participants in health science research who are non-native English speakers.

**METHOD:** The analysis is based on an audio file from one representative participant in a qualitative study involving 20 participants. It compares transcripts generated by a professional audio transcriptionist with those produced by two AI platforms, Otter.ai and Avidnote.

**RESULTS:** Findings reveal that while AI transcription has advantages in speed and cost-effectiveness, it can struggle with speaker differentiation and punctuation accuracy, necessitating manual review. Both platforms faced challenges with cultural terminology and accented speech, but Avidnote showed better performance in word recognition and comprehension. Limitations were primarily in the transcription of te reo Māori.

**CONCLUSION:** The study highlights the critical role of culturally competent researchers in reviewing transcripts to ensure accuracy and clarity. These findings contribute to a deeper understanding of the benefits and limitations of AI transcription tools in qualitative health research, especially when working with linguistically and culturally diverse populations.

Transcription is an important, but often under-acknowledged, aspect of both qualitative research and clinical documentation. In research contexts, transcribing interview data from audio files is a fundamental step for generating high-quality transcripts for qualitative analysis.<sup>1,2</sup> It enables researchers to develop a deep familiarity with the data and facilitates the methodological and theoretical interpretation of spoken language.<sup>2</sup> Transcription is widely used across various academic, applied research and professional practice fields and is also necessary for natural language processing applications such as text summarisation, document clustering, question answering, essay grading and more.<sup>3</sup>

Because spoken language differs structurally from written text, when oral communication is transcribed into written form, readers interpret it based on their understanding of written language. Since the relationships between language and meaning are inherently contextual, rhetorical and constructed, these principles also apply to transcriptions of spoken discourse.<sup>1,4</sup> Transcripts should, therefore, be viewed as contextual constructions rather than objective representations of reality, acknowledging the limitations of capturing all aspects of a speech event.<sup>5</sup>

In qualitative research, transcription is not just a technical step but part of the analytic process,<sup>6</sup> so accuracy and transparency are essential to maintain the trustworthiness of the data. Errors such as missing or incorrect words can significantly alter the meaning and compromise the validity of findings.<sup>7</sup> Different transcription strategies may be used for specific research purposes to capture intonation, false starts, overlaps, speaker positioning and turn-taking.<sup>2</sup> Standard punctuation and spelling also improve the ease of reading by eliminating dysfluencies and accidental starts. Maintaining transcription quality, therefore, is critical for the rigour and reliability of qualitative research.<sup>8</sup>

Some studies have suggested that a combination of manual and automated transcription methods may be the most effective approach, capitalising on the strengths of each to produce the most accurate and comprehensive transcripts.<sup>9,10</sup> With the rapid expansion of artificial intelligence (AI), transcription tools have evolved to support both research and healthcare applications.

AI-driven speech recognition now plays a pivotal role in medical transcription, allowing clinicians to efficiently dictate patient notes and medical records, improving productivity, allowing

for faster and more accurate record-keeping and reducing the risk of manual errors. The ability of speech recognition systems to adapt to medical jargon and nuances makes them increasingly important tools in healthcare documentation.<sup>11</sup> These platforms integrate core technological components, including automatic speech recognition (ASR) engines, natural language processing (NLP), algorithms and user interface systems, although they differ in their underlying architecture, training datasets and performance across diverse linguistic contexts.<sup>12</sup>

Each platform employs distinct methodological approaches to speech recognition: some use proprietary ASR models trained on domain-specific datasets, others support cloud-based application programming interfaces (APIs) from major technology providers, while newer systems incorporate transformer-based architectures and large language models (LLMs) to enhance contextual understanding and accuracy.<sup>12</sup>

In healthcare settings, specialised medical transcription software is often incorporated into patient management packages, such as *indici*,<sup>13</sup> and these are now becoming widely used in general practices across New Zealand. These platforms offer integrated clinical documentation features tailored to the needs of primary care providers, including voice-to-text capabilities.

As AI transcription tools become increasingly popular in research and clinical contexts, understanding their capabilities, limitations and implications is increasingly important. Despite these advantages, human review is often necessary to ensure clinical safety,<sup>11</sup> especially when working with clients with accented English, different cultural expressions or specialised terminology. Given the critical importance of transcription quality for interpretation of data, there is a paucity of literature examining the accuracy, efficiency, reliability and usability of transcripts generated by AI compared with those produced manually by human transcribers.

The expanding field of AI-based transcription technologies includes a diverse range of platforms, ranging from research-oriented tools like *Avidnote*, to general-purpose transcription tools such as *Otter.ai* and *Fireflies.ai*, as well as enterprise-level solutions like *Sonix* and *TurboScribe*. Many of these platforms use advanced speech recognition models to transcribe audio, including in multiple languages. One such model is *OpenAI Whisper*,<sup>14</sup> which is open source and widely used behind the scenes in platforms such as *TurboScribe*

and the *ChatGPT* mobile app. While *OpenAI Whisper* performs the transcription, other tools may use LLMs for downstream tasks such as summarisation or note generation. Similarly, enterprise services like *Microsoft Azure AI Speech* and *Google Cloud Speech-to-Text* are widely used in professional settings, employing their own proprietary speech recognition technologies that support transcription in many languages.<sup>15,16</sup> Given the challenges in accurately transcribing accented English, multilingual speech and culturally specific expressions, we identified a need for a more systematic examination of different tools using representative audio files from target populations.

After an initial assessment of widely available AI transcription tools (see Table 1), *Avidnote* and *Otter.ai* were selected for comparative analysis. This selection was informed by practical relevance and ethical considerations. *Avidnote* was chosen for its research focus, multilingual capability and explicit commitment to ethical data handling practices, specifically compliance with the European Union's General Data Protection Regulation (GDPR). *Otter.ai*, although more limited in its language capabilities, was included due to its widespread use within our university context. The professional subscription versions of both platforms were used in this study, as these provide better data security and privacy compared with free versions.

Table 1 summarises some common AI tools used for transcription, with a particular focus on their ability to transcribe audio in multiple languages. While many tools perform well in English, their accuracy with non-English words and multilingual speech varies significantly.

## Research context

This study emerged from a broader qualitative research project examining dual relationship roles for Muslim professionals who provided services to a community impacted by a terrorist attack on two mosques in Christchurch, New Zealand. Although interviews were conducted in English, participants frequently used Islamic and Arabic terminology, and many spoke English with an accent as it was their second language. An audio transcriptionist was initially engaged and provided with contextual information and a list of commonly used terms. However, she reported difficulty understanding accented English and unfamiliar terminology, which prompted the trial

**Table 1:** Common AI tools for transcription.

AI tool	Primary function	Languages supported	Standalone tool
Alrite AI <a href="https://alrite.io/ai/">https://alrite.io/ai/</a>	Transcription, subtitle generation	Major Indo-European languages	Yes
Avidnote <a href="https://avidnote.com/">https://avidnote.com/</a>	Transcription, note taking, qualitative analysis support	Multiple languages	Yes
Beey AI <a href="https://www.beey.io/en/">https://www.beey.io/en/</a>	Transcription, captions for media content	Major European languages	Yes
Cockatoo <a href="https://www.cockatoo.com/">https://www.cockatoo.com/</a>	Transcription, meeting summarisation	Mostly European & Asian languages	Yes
Deciphrr <a href="https://www.deciphrr.ai/">https://www.deciphrr.ai/</a>	Podcast transcription, summarisation	English	Yes
Fireflies.ai <a href="https://fireflies.ai/">https://fireflies.ai/</a>	Transcription, meeting summarisation	Mostly English	Yes
Otter.ai <a href="https://otter.ai/">https://otter.ai/</a>	Transcription (real time), summarisation, note taking	English	Yes
Sonix <a href="https://sonix.ai/">https://sonix.ai/</a>	Transcription, translation subtitles	Major European & selected Asian languages	Yes
TurboScribe <a href="https://turboscribe.ai/">https://turboscribe.ai/</a>	Transcription (Whisper)	Multiple languages	Yes
Whisper (OpenAI) <a href="https://openai.com/index/whisper">https://openai.com/index/whisper</a>	Automatic speech recognition (used in other platforms like TurboScribe & ChatGPT mobile app)	Multiple languages	No (back-end model integrated into other platforms)
Microsoft Azure AI speech services <a href="https://learn.microsoft.com/en-us/azure/ai-services/speech-service/overview">https://learn.microsoft.com/en-us/azure/ai-services/speech-service/overview</a>	Transcription, summarisation	Multiple languages & dialects	No (integrated in Microsoft 365 apps)
Google Cloud Speech-to-text <a href="https://cloud.google.com/speech-to-text/docs">https://cloud.google.com/speech-to-text/docs</a>	Transcription, speaker separation	Multiple languages & dialects	Yes (Google Cloud API*)

\*API = application programming interface.

Note: Descriptions of platform capabilities are based on publicly available documentation and may vary depending on subscription level, integration settings or usage context.

of Otter.ai and Avidnote to assess their potential for research with multicultural participants.

The primary aim of this study was to evaluate discrepancies between AI-generated and human-generated transcripts, while also exploring how cultural comprehension may help researchers interpret nuanced contexts within the transcripts. This evaluation focussed on the semantic relationships among words, sentences and concepts, using semantic similarity metrics.<sup>3</sup> These metrics have rarely been used to assess the fundamental differences between AI-generated and manually processed transcripts. In this paper, we report sentence-to-sentence semantic similarity analysis to evaluate the potential qualities of both AI and manual transcription methodologies.

This analysis was guided by the following research questions:

1. Does the involvement of culturally competent researchers as transcribers impact the accuracy and quality in qualitative research transcripts?
2. What are the potential challenges and benefits of using AI transcription tools compared with human transcribers to accurately capture cultural nuances and context?
3. How can researchers effectively integrate AI-generated transcripts into their qualitative research process to maintain rigour and accuracy?

## Methodology

The comparative analysis utilised an audio file from a single participant, coded as DR06 to ensure anonymity. This participant was part of a larger sample, and the issues discussed in the interview, as well as the transcription concerns identified by the human transcriber, were representative of the broader dataset. The participant provided informed consent for their data to be used in this analysis, including further transcription by AI systems. The audio file was recorded during a semi-structured interview aimed at exploring the participant's experiences and perspectives on their professional role working in a Muslim community after the terrorist attack. The interview lasted 21 minutes and was conducted in a private setting to ensure confidentiality and encourage candid responses. The audio file size was 216MB and was recorded as part of a study approved by the University of Otago Human Ethics Committee

(Health), approval number 22/153. Prior to data collection, participants received comprehensive information about the study and provided written consent.

### Manual transcription

The audio file was first transcribed manually by an experienced, professional transcriptionist. She produced a verbatim transcript, capturing spoken words, fillers and pauses.

### AI transcription

Two AI platforms, Otter.ai and Avidnote, were used to generate automated transcriptions of the same audio file. Both platforms are designed for academic and research purposes, with features that include transcription, annotation and organisation of notes.

- **Otter.ai:** the audio file was uploaded directly to the program, and the default settings were used to generate the transcript.
- **Avidnote:** this AI-driven transcription program is designed for academic use. The DR06 file was uploaded and transcribed using standard settings. Avidnote supports files of up to 500MB; however, larger files require splitting into smaller segments using additional software.

### Culturally responsive review

Two research team members, with expertise in the subject matter and familiarity with the participants' cultural context, thoroughly reviewed all transcripts, comparing them against the audio file. They noted whether cultural nuances, idiomatic expressions and context-specific terminologies were accurately captured. The annotated and reviewed transcript produced through this process was then used as the reference standard for subsequent comparisons between manual and AI-generated transcriptions.<sup>17</sup>

### Comparative analysis

A semantic similarity approach was used to compare the different transcript versions and identify the strengths and limitations of each method.

### Lexical and morphological differences

The primary focus of the comparative analysis was on lexical and morphological differences. Lexical differences refer to variations in word choice, while morphological differences relate to

the structure and form of the words. This analysis involved several steps:

1. Initial review: each transcript was reviewed independently to identify obvious differences.
2. Lexical analysis: a detailed comparison was conducted to identify differences in vocabulary used across the transcripts. This included noting different words or phrases used to convey the same meaning, as well as omissions or additions of words.
3. Morphological analysis: comparing the structure of words and their grammatical forms. Special attention was paid to foreign language words and expressions.
4. Quantitative measures: the number of lexical and morphological discrepancies in each transcript was counted, providing a numerical basis for comparing the accuracy of the different transcriptions.

## Findings

Distinct patterns of errors were observed across the different transcription methods. The distribution and types of errors for each

transcription method are detailed in Table 2.

For the manual transcription, the main sources of error stemmed from the transcriptionist's inability to accurately capture foreign terminology and non-English words, combined with her difficulty understanding diverse accents. A total of 17 errors were recorded. In contrast, for Otter.ai, a total of 84 errors were identified. The majority of these were morphological and lexical, with a significant number also related to inaccurate punctuation. Although Otter.ai performed reasonably well in recognising and transcribing words, it struggled with the finer details of language structure and punctuation. It also had trouble with foreign terms. Avidnote, on the other hand, produced the most accurate transcript, with 16 errors detected. These were predominantly lexical and could likely be attributed to the software's challenges in accurately processing accented pronunciation.

## Challenges with non-English terminology and accents

The professional transcriber faced challenges in identifying cultural words and understanding certain accents. Words such as *Masjid*, *Qadar* and *Shahada*, which are commonly used in this

**Table 2:** A list of errors (morphological,<sup>1</sup> lexical,<sup>2</sup> non-English & punctuation) compared with the reference version.

Manual transcription	Otter.ai-generated transcript	Avidnote-generated transcript
Errors found: 17	Errors found: 84	Errors found: 16
Types of error: Morphological = 2 Lexical = 1 Non-English words = 11 Punctuation errors = 3	Types of error: Morphological = 35 Lexical = 21 Non-English words = 5 Punctuation errors = 23	Types of error: Morphological = 4 Lexical = 7 Non-English words = 3 Punctuation errors = 2
Examples: Non-English words <i>Masjid</i> or <i>Madrasa</i> were missed.	Examples: Morphological, e.g., "...he was present in the <b>material</b> know, the whole thing..." instead of "...he was present <b>and he</b> know the whole thing..." Lexical, e.g., "...it was an evolving <b>road</b> " instead of "...it was an evolving <b>role</b> " Non-English words, e.g., "...see them in the <b>dresser</b> " instead of "...see them in the <b>Madrasa</b> "	Examples: Morphological, e.g., "...offering them support <b>in like</b> letting them..." instead of "...offering them support <b>and</b> letting them..." Lexical, e.g., " <b>stuff</b> " instead of " <b>staff</b> "; " <b>rules</b> " instead of " <b>roles</b> "; " <b>dual rooms</b> " instead of " <b>dual roles</b> "

<sup>1</sup> A type of error that occurs when the rules of word formation in a language are not followed correctly.

<sup>2</sup> A type of error that is associated with the incorrect use of vocabulary and collocation.

community, were not accurately transcribed (see Table 3). She also had difficulty with the accent of DR06, who speaks English as a second language, failing to capture phrases like “*I actually didn’t really have a profession, even though I’m an MBA in...*”, the transcriber failed to perceive “an MBA in”. As a native English speaker, she had limited familiarity with diverse accents. External factors like audio quality may have also impacted her comprehension. In contrast, the multicultural research team had no trouble understanding the participant.

Although Avidnote, and to a lesser extent Otter.ai, correctly identified many cultural and religious terms, they both struggled with te

reo Māori words. The manual transcriptionist accurately captured the Māori content, despite not being fluent in the language. Though the speaker is not Māori, they, like many New Zealanders, have a basic understanding of te reo Māori terminology. This suggests that the broader implications of the transcriber’s linguistic background and competency are important to consider.

### Punctuation errors

Another issue relates to punctuation accuracy, particularly for Otter.ai, which had 23 errors detected. Punctuation often relies on understanding the context of a sentence,<sup>18</sup> and while AI may excel in transcribing speech, it can have difficulty in

**Table 3:** A semantic comparison between AI-generated and manually processed transcriptions.

Professional transcriber	Otter.ai-generated	Avidnote-generated	Culturally informed researcher
1. I was volunteering with the <b>01:32</b> Council with Sister...	I was volunteering with the <b>nice Asherah council with system...</b>	I was volunteering with the (correctly identified organisation) <b>Council</b> with sister...	I was volunteering with (correctly identified organisation) <b>Council</b> with sister...
2. And as you know the ... community suffered the largest numbers of <b>02:20</b> .	And as you know, the, the ... community suffered the largest numbers of <b>Shahada</b> .	And as you know, the, the ... community suffered the largest numbers of <b>Shahada</b> .	And as you know, the, the ... community suffered the largest numbers of <b>Shahada</b> .
3. ...when the teams of community support workers <b>at Purapura Whetu</b> was being established...	...when the teams of community support workers, <b>a political figure</b> was being established...	...when the team of community support workers <b>at Poro Poro 52</b> was being established...	...when the teams of community support workers <b>at Purapura Whetu</b> was being established...
4. I’m not a very religious person but I’m very spiritual and I believe in all <b>07:48</b> and destiny.	I’m not a very religious person, but I’m very spiritual, and I believe <b>of all in all, color</b> and destiny	I’m not a very religious person, but I’m very spiritual, and I believe in all <b>Qadar</b> and destiny.	I’m not a very religious person, but I’m very spiritual, and I believe in all <b>Qadr</b> and destiny.
5. He was present <b>06:13</b> and the whole thing and he came back.	He was present in the <b>material know</b> , the whole thing. And he came back.	He was present <b>in Masjid Al Noor</b> , the whole thing, and he came back.	He was present <b>in Masjid Al Noor</b> ; the whole thing and he came back.
6. I would do it again because I feel like <b>Allah’s 12:27</b> has bigger plans, I am just one small piece in that.	I will do it again. Because I feel like <b>Allah Subhanaw taala</b> has bigger plans.	I will do it again, because I feel like <b>Allah subhanahu wa ta’ala</b> has bigger plans.	I will do it again because <b>Allah Subhanahu Wa Ta’ala</b> has bigger plans.
7. I actually didn’t really have a <b>profession</b> , even though I’m <b>4:36</b> ...	I actually didn’t really have a <b>question</b> , even though I’m <b>an MBA</b> in...	I actually didn’t really have a <b>profession</b> , even though I’m <b>an MBA</b> in...	I actually didn’t really have a <b>profession</b> , even though I’m <b>an MBA</b> in...

accurately interpreting the context and appropriately inserting punctuation marks. Like any technology, AI may face challenges in accurately capturing the nuances of spoken language, including pauses and intonations that indicate where punctuation should be placed.

### Cost and time savings

Manual transcription generally involves a turnaround time of 2 to 3 days. In contrast, Otter.ai and Avidnote provided significantly faster processing times, completing transcriptions in 7.56 minutes and 3.42 minutes, respectively. In addition to improving workflow efficiency, they also offer potential cost savings, with the annual subscription fees for these AI-transcription software services being approximately equivalent to the cost of a single manual transcription.

### Discussion

This study investigates the advantages and limitations of two AI-based transcription platforms, Otter.ai and Avidnote, compared with traditional manual transcription methods. The findings suggest that while AI transcription tools have the potential to significantly enhance workflows in research and healthcare settings, they are not without limitations. Key benefits of these AI platforms include significant time and cost savings, as well as the ability of Avidnote to process files in multiple languages. However, several challenges were identified. For instance, although Otter.ai has a user-friendly interface with a streamlined process for uploading large audio files, the transcript it generated contained numerous morphological, lexical and punctuation errors. Avidnote, on the other hand, is limited to processing audio files no larger than 500MB, necessitating the manual segmentation of larger files before uploading. This step introduces additional technical complexity and software requirements.

Both AI platforms encountered difficulties in accurately transcribing New Zealand Māori words, highlighting a potential bias in transcription algorithms towards more widely spoken languages. Avidnote demonstrated superior accuracy in transcribing multilingual content, a task that presented challenges for the manual transcriptionist, especially with non-English terminology and diverse accents.

Despite the rapid transcription capability of AI platforms, human oversight remains crucial. Conversations are inherently complex, built upon

multiple layers of context, shared understandings and cultural assumptions.<sup>2</sup> Staff with cultural expertise or insider perspectives are better equipped to identify and interpret these nuances, which may be overlooked by professional transcribers or AI-based systems, and this can significantly enhance transcription quality.<sup>8</sup> In clinical practice, this suggests that healthcare providers should prioritise involving professionals who understand the cultural backgrounds of their patients. Such an approach can improve communication, build trust and ensure that healthcare interventions are appropriately tailored to meet the needs of linguistically diverse groups.

While AI transcription tools can improve efficiency, it is essential to thoroughly review and edit AI-generated transcripts to preserve the rigour and trustworthiness of research. Existing literature indicates that AI-generated transcriptions are susceptible to bias, which highlights the necessity of human intervention to maintain research integrity.<sup>19</sup> Given that preliminary verbatim transcripts require verification, revision and elaboration by researchers,<sup>2</sup> it is crucial that AI-generated transcripts also undergo meticulous examination and refinement to ensure their accuracy. Our findings recommend a hybrid approach, where AI is used to rapidly generate initial transcripts, followed by thorough review and refinement by culturally competent researchers. Particular attention should be given to accurately transcribing and interpreting linguistic nuances, especially for individuals with accents or who use nonstandard language. Collaboration with language experts or consultation with interview participants, when necessary, will effectively preserve the richness and authenticity of participants' voices, thereby improving the overall quality of the transcripts.

### Implications of transcription quality in clinical practice and healthcare

The findings from this study have significant implications for clinical practice and healthcare. AI transcription tools, while offering substantial benefits for efficiency and cost-effectiveness, do have some limitations, especially in accurately transcribing culturally specific terminology and accented speech. In healthcare settings, where understanding patient narratives is crucial for diagnosis and treatment, inaccuracies in transcripts may lead to misinterpretations of patient needs and cultural contexts, potentially resulting in inappropriate clinical decisions or harm.

These issues could significantly affect the quality of healthcare and clinical practice provided to culturally and linguistically diverse patients.<sup>20,21</sup>

### Limitations of this research study

This research is limited by the analysis of a single audio file from one participant, which may not fully represent the broader dataset or the diversity of linguistic and cultural experiences across all contexts or populations, nor does it explore the full range of factors that could influence transcription accuracy, such as the specific training of human transcribers. The comparison was also restricted to two AI transcription platforms, Otter.ai and Avidnote, as well as manual transcription, which limits the generalisability of the findings. While the findings offer useful insights into the capabilities and limitations of AI transcription tools, it does not include the full range of technologies available, especially those specifically designed for clinical use.

Future studies should consider evaluating additional AI transcription tools, including healthcare-specific platforms that are becoming increasingly common in New Zealand. Expanding the sample to include multiple participants with varied linguistic backgrounds would also provide a more comprehensive understanding of transcription accuracy and cultural sensitivity in different contexts. Future research could also explore how the training, experience and cultural competence of human transcribers influence transcription accuracy.

### Conclusion

This study compared transcriptions generated by two AI platforms, Avidnote and Otter.ai, with those produced by a professional human transcriptionist. Culturally knowledgeable researchers reviewed the original transcripts and created a reference transcript that served as the benchmark for evaluating the others.

Overall, AI transcription tools offer greater

speed and cost-effectiveness than manual methods. Avidnote successfully transcribed an interview conducted in English with a non-English accent and accurately captured Arabic phrases used within this context. However, its inability to process files larger than 500MB necessitated manual segmentation. Otter.ai was also fast, cost-effective and capable of managing large audio files, but it was less accurate with non-English terminology and introduced multiple other errors, potentially requiring substantial correction time. Manual transcription, while slower and more costly, benefits from the cultural knowledge and contextual understanding of the transcriber, which can be critical for accurately capturing nuanced or culturally specific language. In this study, the human transcriber encountered challenges in accurately capturing certain elements, such as the participant's accent and Arabic terminology, although Māori terminology was transcribed successfully.

Despite the efficiency gains offered by AI-powered transcription tools, thorough review by culturally competent researchers remains necessary to ensure the accuracy and cultural relevance of the final transcripts. These findings have implications for clinical practice, particularly when documenting consultations from individuals from diverse cultural backgrounds or with varying levels of English proficiency. Transcripts, or any summaries produced, could be compromised if there are errors and should therefore be carefully reviewed.

The findings indicate that while AI-powered tools can expedite the transcription process, they still lack the cultural sensitivity and nuanced understanding necessary to produce high-quality transcripts, particularly within the diverse cultural contexts encountered in clinical practice and health science research. A hybrid approach, using AI for initial transcription followed by culturally informed human review, is recommended to ensure accuracy and contextual relevance.

**COMPETING INTERESTS**

Canterbury Medical Research Foundation (CMRF) Major Project Grant (Sulaiman-Hill MPG 2022) funding for dual relationship study.

**ACKNOWLEDGEMENTS**

We sincerely thank Professor Richard Porter for his invaluable support and critical feedback.

**AUTHOR INFORMATION**

Dr SM Akramul Kabir: Research Fellow, Department of Psychological Medicine, University of Otago Christchurch.

Fareeha Ali: Assistant Research Fellow & PhD Candidate, Department of Psychological Medicine, University of Otago Christchurch.

Dr Ruqayya Sulaiman-Hill: Senior Research Fellow, Department of Psychological Medicine, University of Otago Christchurch.

**CORRESPONDING AUTHOR**

Dr SM Akramul Kabir: Research Fellow, Department of Psychological Medicine, University of Otago Christchurch, PO Box 4345, Christchurch 8140. E: sm.akramul.kabir@otago.ac.nz

**URL**

<https://nzmj.org.nz/journal/vol-138-no-1625/a-comparative-assessment-of-ai-and-manual-transcription-quality-in-health-data-insights-from-field-observations>

**REFERENCES**

1. Kvale S. An introduction to qualitative research interviewing. Sage; 1996.
2. Lapadat JC. Problematizing transcription: Purpose, paradigm and quality. *Int J Soc Res Methodol* . 2000;3(3):203-19.
3. Oussalah M, Mohamed M. Knowledge-based sentence semantic similarity: algebraical properties. *PROG ARTIF INTELL*. 2022;11:43-63.
4. Smagorinsky P. If meaning is constructed, what is it made from? Toward a cultural theory of reading. *Rev Educ Res*. 2001;71(1):133-169.
5. Poland BD. Transcription quality. In: Gubrium J, Holstein J, eds. *Handbook of interview research: context and method*. Sage; 2001:629-649.
6. Skukauskaite A. Transparency in transcribing: Making visible theoretical bases impacting knowledge construction from open-ended interview records. *FQS*. 2012;3(1):1-32.
7. Poland BD. Transcription quality as an aspect of rigor in qualitative research. *Qual Inq*. 1995;1(3):290-310.
8. Witcher CSG. Negotiating transcription as a relative insider: Implications for rigor. *Int J Qual*. 2010;9(2):122-132.
9. Clausen AS. The individually focused interview: Methodological quality without transcription of audio recordings. *TQR*. 2012;17(19):1-17.
10. Welsh E. Using NVivo in the qualitative data analysis process. *FQS*. 2002;3(2):26-34.
11. Sarkar U, Bates DW. Using Artificial Intelligence to Improve Primary Care for Patients and Clinicians. *JAMA Intern Med*. 2024;184(4):343-344. doi:10.1001/jamainternmed.2023.7965
12. Mehta D, Upadhyay R, Jariwala K. Integrating Speech Recognition and NLP for Efficient Transcription Solutions. *IJSRCSEIT*. 2025;11(1):1089-1096. doi:10.32628/cseit2526479
13. Valentia Technologies (NZ) Limited. Indici: Cloud-based electronic health record platform [Internet]. Valentia Technologies; 2023 [cited 2025 Aug 12]. Available from: <https://www.indici.co.nz>
14. Andreyev A. Quantization for OpenAI's Whisper Models: A Comparative Analysis. arXiv preprint. 2025. doi:10.48550/arXiv.2503.09905
15. Google Cloud. Speech-to-Text documentation [Internet]. 2025 [cited 2025 Sep 25]. Available from: <https://cloud.google.com/speech-to-text/docs/>
16. Microsoft. Speech to text overview – Azure AI services [Internet]. 2025 [cited 2025 Sep 25]. Available from: <https://learn.microsoft.com/en-us/azure/ai-services/speech-service/speech-to-text>
17. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol*. 2006;3(2):77-101.
18. Moore N. What's the point? The role of punctuation in realising information structure in written English. *Functional Linguist*. 2016;3(6):1-23.
19. Eftekhari H. Transcribing in the digital age: qualitative research practice utilizing intelligent speech recognition technology. *Eur J Cardiovasc Nurs*. 2024;23(5):553-560.
20. McMullin C. Transcription and Qualitative Methods: Implications for Third Sector Research. *Voluntas*. 2023;34(1):140-153. doi: 10.1007/s11266-021-00400-3.
21. Mero-Jaffe I. 'Is that what I said?' Interview transcript approval by participants: An aspect of ethics in qualitative research. *Int J Qual*. 2011;10(3):231-247.