International Journal of Education Humanities and Social Science

ISSN: 2582-0745
Vol. 7, No. 04; 2024

DATA COLLECTION PROTOCOL FOR ORAL CORPUS OF MAJOR DEPRESSION DISORDER PATIENT

Salihah Ab Patah[1*], Norliza Jamaluddin[2], Ng Chong Guan[3] and Ermanto[4]
[12]Jabatan Bahasa Melayu dan Kesusasteraan, Fakulti Bahasa dan Komunikasi, Universiti Pendidikan Sultan Idris, 35900, Tg Malim, Perak, Malaysia.
[3]Jabatan Perubatan Psikologi, Fakulti Perubatan, Universiti Malaya, 59100, Kuala Lumpur, Malaysia.
[4]Universitas Negeri Padang, Indonesia

https://doi.org/10.54922/IJEHSS.2024.0752

ABSTRACT
The collecting of oral data involving mental health patients in hospitals demands research ethics approval to mitigate risks or harm towards participants of the study. The appointment of psychiatric experts as clinical supervisors is imperative to facilitate the process of data collection through interviews, particularly with patients diagnosed with Major Depressive Disorder (MDD). The process of obtaining permission to access hospitals and interview patients involves research ethics approval from the Research Management and Innovation Centre, Sultan Idris Education University (RMIC-UPSI), and the Research Ethics Committee of the University of Malaya Medical Centre (UMMC-MREC). A total of 35 MDD patients were interviewed according to the study protocol and interview guidelines. The interview protocol is also subject to the Prevention and Control of Infectious Diseases Act 1988 (Act 342) due to the Covid-19 pandemic in Malaysia. The transcribed interview data can amount to 65,894 words when extracted into WordSmith 8.0 software. Overall, this data collection process encompasses, appointment of psychiatric experts as clinical supervisors, research ethics approval and field data collection protocols.

Keywords: Protocol, Patient, Depression, Corpus Data.

1. INTRODUCTION
In the collection of data to form an oral corpus, each researcher is responsible for ensuring that ethical issues and study procedures are conducted and reported transparently. When collecting data involving human behavior and social life, there are ethical principles that need to be adhered to and implemented during the planning and execution of the study (Creswell, 2018). In qualitative data collection, the interview method is the most common and frequently used approach (Frels et al., 2013 & Friginal et al., 2014). Oral data collection based on demographic parameters requires specific procedures and protocols (M.F. Zaini et al., 2022). Adherence to standardized procedures by researchers can alleviate risks or harm to participants (Denscombe, 2017).

In this study, clinical data collection was conducted through linguistic interview methods. For linguistic researchers, exposure to fieldwork in hospitals such as healthcare personnel is often limited. This lack of exposure hinders linguistic studies in the medical field despite their potentially more authentic data (Demjén, 2020). Therefore, interdisciplinary studies like this require collaboration between academic professionals and government institutions (Delgadillo, 2016). As this study focuses on mental patients, researchers require the expertise of psychiatrists to facilitate data collection in hospitals. In other words, psychiatrists have access to patients as they are trained

ISSN: 2582-0745
Vol. 7, No. 04; 2024

professionals capable of assessing, classifying, and treating emotional and behavioral issues through physical examinations, patient history, and laboratory findings (Abd El-Hay, 2018).

Among the ethics to be observed when conducting data collection, especially involving mental patients, is obtaining permission to enter hospitals and interview patients. This permission process involves two parties: the Research Management and Innovation Centre, Sultan Idris Education University (RMIC-UPSI), and the Research Ethics Committee of the University of Malaya Medical Centre (UMMC-MREC). Interviews conducted with 35 patients at UMMC followed study protocols and interview guidelines. Given that data collection occurred during the Covid-19 Movement Control Order (MCO), interview protocols also adhere to the Prevention and Control of Infectious Diseases Act 1988 [4]. In this study, interview data transcripts were extracted into oral corpus data. Developing oral corpus data according to this purposive sampling is necessary for qualitative research. Figure 1 illustrates the data collection protocol for the oral corpus of Major Depressive Disorder (MDD) patients.

2. DESCRIPTION OF PROTOCOL

Protocol for obtaining research ethics approval.

For researchers to access hospitals and conduct interviews with patients at the University of Malaya Medical Centre (UMMC), research ethics approval from RMIC-UPSI and UMMC-MREC is required. The appointment of psychiatric experts as clinical supervisors is also undertaken to facilitate the process of obtaining research ethics approval. This appointment process involves the Faculty of Languages and Communication (FBK) and the Institute of Graduate Studies (IPS). Additionally, the services of psychiatric experts are required prior to data collection. Table 1 illustrates the roles of psychiatric experts in data collection, while Table 2 outlines the four steps to obtain research ethics approval.

ISSN: 2582-0745
Vol. 7, No. 04; 2024

Table 1. The role of the psychiatrist

	The steps	Action	Methodology	Action
The role of the psychiatrist	Playing a role in linguistic research	Collaborate with linguists.	Agree to be appointed as a clinical supervisor. Give resumes to researchers to send to FBK. Received appointment as clinical supervisor from IPS. Discussing with researchers online. Evaluate the study protocol and interview questions. Set 4 patient criteria, namely MDD patients who have been diagnosed, aged 18 years and above, of Malay descent and fluent in Malay. Recommended interview time, which is between 45 minutes to 1 hour. Represent researchers to present research at UMMC-MREC online. Suggest patients who can be interviewed on the day of data collection.	Researchers, psychiatrists, FBK, IPS and UMMC-MREC.

ISSN: 2582-0745
Vol. 7, No. 04; 2024

Table 2. Four steps in obtaining research ethics approval.

	The steps	Action	Methodology	Involvement
Applying research ethics	Apply for research ethics from RMIC-UPSI	Refer to the RMIC-UPSI website as a guide.	Submit the human research ethics form. Submit research and interview protocols. Improving study and interview protocols after being rejected the first time. Make an online research presentation. Deliver a letter of appointment of psychiatrist as clinical supervisor from IPS after two weeks of presentation. Get approval to conduct interviews for a year.	researcher, RMIC-UPSI and IPS
	Get the services of a psychiatrist	Contact psychiatrists at UMMC via email.	Submitting the study protocol and interview guidelines. Scheduling appointments and discussing via Zoom software. Obtaining confirmation from psychiatric experts regarding interview questions and the four criteria for patients to be interviewed.	researchers and psychiatrists
	Appoint a psychiatrist as a clinical supervisor	Referring to the FBK website as a guide.	Completing the clinical supervisor application form. Submitting the resume of the psychiatric expert to FBK. Obtaining certification from the psychiatric expert as a clinical supervisor.	researchers and psychiatrists, FBK and IPS

ISSN: 2582-0745
Vol. 7, No. 04; 2024

		Obtaining a letter of appointment for the psychiatric expert as a clinical supervisor from IPS.	
Apply for research ethics from UMMC-MREC.	Refer to the UMMC-MREC website as a guide.	Submitting the human research ethics application form. Submitting the study protocol and interview guidelines. Revising the study protocol and interview guidelines after the first rejection. Delegating a psychiatric expert to present the research proposal online. Obtaining approval to conduct interviews for one year.	researchers, psychiatrists and UMMC-MREC

Protocol for field data collection

Following the research ethics approval from RMIC-UPSI and UMMC-MREC, data collection commences in accordance with the study protocol and interview guidelines. On the day of data collection, psychiatric and medical services are required to identify patients based on the four criteria specified in the study protocol. Among the ethical issues communicated to patients is the assurance of confidentiality regarding patient data. In this regard, the true identities of patients are kept confidential, and any sensitive information obtained is anonymized, processed, and paraphrased without compromising the research outcomes. In other words, collected data must be managed securely and treated as confidential. During the data collection period, compliance with safety regulations is enforced due to the Covid-19 pandemic in Malaysia. Table 3 illustrates three steps for data collection at UMMC, while Table 4 outlines procedures for ensuring safety from the Covid-19 outbreak.

ISSN: 2582-0745
Vol. 7, No. 04; 2024

Table 3. Three steps to collect data in the field

	The steps	Action	Methodology	Engagement
collect data in the field	Get the services of a psychiatrist	Discuss with a psychiatrist.	Requesting the expertise of a psychiatric expert to recommend patients according to the four criteria established in the study protocol.	researchers and psychiatrists
	Get the services of a medical officer	Discuss with the medical staff.	Introducing oneself and explaining the purpose of the research. Requesting medical personnel to identify patients based on the four criteria established in the study protocol. Requesting the patient's medical card as reference. Seeking permission from healthcare personnel to interview patients in the treatment room. Requesting medical personnel to sign the informed consent form as witnesses.	researchers and medical experts
	Get the MDD patient's cooperation	Recruit patients.	Calling out the patient's name in the waiting area. Introducing oneself. Explaining the research objectives and stating the importance of their involvement. Presenting the research information sheet. Allowing time for the patient to express consent. Seeking the patient's permission to conduct the interview in the treatment room after obtaining oral consent.	researchers and MDD patients

ISSN: 2582-0745
Vol. 7, No. 04; 2024

			Requesting the patient to voluntarily sign the informed consent form before commencing the interview. Conducting the patient interview for approximately 45 minutes to one hour. Providing compensation to the patient after the interview concludes.	

Table 4. Procedures to keep safe from the Covid-19 epidemic.

	The Steps	Action	Methodology	Engagement
Compliance with the Prevention and Control of Infectious Diseases Act 1988 (Act 342) [Akta Pencegahan dan Pengawalan Penyakit Berjangkit 1988.].	Understand the ethics of conducting face-to-face research.	Referring to the website of the Ministry of Health Malaysia.	Wear a face mask while in the hospital. Maintain social distancing with patients during interviews. Do not make any skin contact such as greeting while in the hospital. Wash your hands with disinfectant every time you conduct an interview.	MDD researchers and patients

3. PROCEDURE FOR CONSTRUCTING A DO-IT-YOURSELF (DIY) CORPUS

Qualitative data collection based on computational corpus data involves very small sample sizes (Friginal, 2014). Moreover, the size of oral corpus data is smaller compared to written corpus data due to ethical considerations, copyright issues, and data transcription processes (Adolphs et al, 2010). Additionally, the construction of a corpus depends on the research objectives and questions Friginal et al., 2014 & Norliza, 2015). To estimate the number of study participants to be interviewed in the actual study, a pilot study is conducted. This is stated because the size of the corpus data will affect the number of required respondents.

The results of the pilot study indicate that a patient can speak only between 1200 to 2000 words. This demonstrates that depressed patients speak slowly and sometimes remain silent for a while, uttering a few or several words only, followed by a low, monotone, and repetitive voice tone (Affizal, 2017). In this study, a total of 35 patients were interviewed. Each interview recording was transcribed verbatim before being extracted into WordSmith 8.0 software. Therefore, the size of the developed oral corpus data is 65,594 words. Demjén (2015) estimates that specific corpus data consisting of written data such as diaries and narratives are adequate between 43,500 to 77,900

ISSN: 2582-0745
Vol. 7, No. 04; 2024

words. Comparing oral corpus data with written data, a size of 65,594 words is considered substantial. Table 5 outlines the procedure for constructing a DIY corpus.

Table 5. The procedure for building your own corpus data.

	The Steps	Action	Methodology	Engagement
Construction of own corpus data	Determining the sampling of oral corpus data.	Estimating the number of occurrences of words for spoken corpus data.	Conducting a literature review. Conducting a pilot study. Determining the average number of words produced by a patient. Transcribing each interview data before conducting the next interview. Uploading the transcribed interview data into WordSmith 8.0 software. Developing an oral corpus data consisting of 65,594 words. Ceasing interviews once the corpus data size is deemed sufficient to address the research questions.	Researchers and patients.

4. CONCLUSION

This article discusses the data collection protocol for a spoken corpus of patients with Major Depressive Disorder (MDD). The data collection protocol involves four steps to obtain research ethics approval, three steps for field data collection, compliance with safety acts, and procedures for building a self-compiled corpus. The study and interview protocols, along with the appointment of a psychiatrist as a clinical supervisor, can serve as guidelines for linguists who wish to use clinical data in their research. Adherence to safety acts throughout data collection indicates that language studies in the medical field can be conducted even during the COVID-19 pandemic. However, data collection is not bound by any safety acts if the country is free from infectious outbreaks. The computational corpus database used in this study reflects that the linguistic data consists of native speakers, namely MDD patients. This demonstrates that the developed corpus data is not fabricated by the researchers but exists naturally, is authentic, and reflects the actual language behavior of depression. Therefore, the size of the spoken corpus developed, totaling 65,594 words, has proven to enhance reliability for use as a language sample in linguistic research.

Protocol validation

*Results and Discussion are **not** part of a methods article.*

Limitations

'None' or 'Not applicable'.

ISSN: 2582-0745
Vol. 7, No. 04; 2024

Credit Author Statement

Salihah Ab Patah: *writer & original idea of writing****, Norliza Jamaluddin:*** *lead supervisor, methodology,* ***Ng Chong Guan:*** *clinical supervisor, psychiatry, interview protocol, validity, data sources,* ***Ermanto****: evaluator & reviewer*

Acknowledgments

This research did not receive any specific grants from funding agencies in the public, commercial or not-for-profit sectors.

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

REFERENCES

[1] Abd El-Hay, Mohamed. (2018). The Psychiatric Interview. 10.4324/9781315148137-3.
[2]Adolphs, S., & Knight, D. (2010). Building a spoken corpus. The Routledge handbook of Corpus Linguistics, 38-52.
[3 Affizal Ahmad. (2017). Kemurungan Mengancam Kehidupan. Kuala Lumpur: Dewan Bahasa
[4] Akta Pencegahan dan Pengawalan Penyakit Berjangkit 1988. Infografik.
https://covid-19.moh.gov.my/faqsop/akta-342
[5] Creswell, J.W. (5th ed.). (2018). Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. Thousand Oaks: CA. Sage Publications.
[6] Delgadillo, Lucy. (2016). Best Practices for Collaboration in Research. Family and Consumer Sciences Research Journal. 45. 5-8. 10.1111/fcsr.12175.
[7] Demjén, Z. (2015). Sylvia Plath and the Language of Affective States: Written Discourse and the Experience of Depression. Bloomsbury Academy. An imprint of Bloomsbury Publishing Plc
[8] Demjén, Z. (2020). Applying Linguistics in Illness and Healthcare Contexts: Contemporary Studies in Linguistics. 10.5040/9781350057685.
[9] Denscombe, Martyn. (2017). The Good Research Guide For Small Scale Social Research Projects (4th ed.). Buckingham: Open University Press.
[10] Dilshad, R. M. & Latif, M. I. (2013). Focus Group Interview as a Tool for Qualitative Research: An Analysis. Pakistan Journal of Social Sciences, 33(1), 191-198.
[11] Frels, R. K. & Onwuegbuzie, A. J. (2013). Administering Quantitative Instruments with Qualitative Interviews: A Mixed Research Approach. Journal of Counselling & Development,
[12] Friginal, E. & A. Hardy, J. (2014). Corpus-Based Sociolinguistics: A Guide for Students. Routledge Taylor & Francis Group. New York and London.
[13] M.F. Zaini, M.M Muhammad, N. Jamaluddin, M.Z.N. Md. Yusoff, N. Hamzah, N.Z. Muhd Zulkifli, M.H Mohd Tahir, S. Pillai (2022). Protocol Methodology for Permission Release in the Construction of a Written Corpus.MethodsX. 9. 101754. 0.1016/j.mex.2022.101754.
[14] Norliza Jamaluddin. (2015). Perlakuan Kata Sifat Berasaskan Data Korpus Berkomputer. Kuala Lumpur: Dewan Bahasa dan Pustaka.