



Developing Ethical Guidelines for the Use of Artificial Intelligence in Speech-Language Pathology Practice in Pakistan: A Mixed-Methods Study

Ammara Manzoor Malik¹, Dr. Syeda Fareesa Hassan Zaidi², Areba Jabeen³ & Jawaria Ambreen⁴

¹Speech and Language Pathologist, MS in Speech and Language Pathology, Riphah International University, Islamabad, Corresponding Author Email: ammaramalik56.am@gmail.com

²Speech and language pathologist/ Physiotherapist, MS-SLP(Riphah International University, Islamabad), DPT(University of Lahore), Email: fareesazaidipt@gmail.com

³BS Behavioral Sciences (Fatima Jinnah Women University), Islamabad MS Speech and language pathology (Riphah International University), Islamabad, Email: arebajabeen96@gmail.com

⁴Speech and Language Pathologist, MS in Speech and Language Pathology, Riphah International University, Lahore Campus, Email: j.ambreen0300@gmail.com

ARTICLE INFO

Article History:

Received:	January	22, 2026
Revised:	March	14, 2026
Accepted:	March	21, 2026
Available Online:	March	28, 2026

Keywords:

artificial intelligence; ethics; speech-language pathology; Pakistan; mixed methods; Delphi; cultural validity; data privacy; professional accountability

Corresponding Author:

Ammara Manzoor Malik

Email:

ammaramalik56.am@gmail.com

ABSTRACT

Artificial intelligence (AI) is revolutionizing healthcare services around the world, such as speech-language pathology (SLP). In Pakistan, SLP is still a developing profession with only a few regulatory bodies, there are no formal codes of ethics for the use of AI in clinical practice. The purpose of this mixed-methods study was to create an ethical framework for the application of AI in Pakistani SLP practice that is grounded in the local context. The aim of this mixed-methods study was to develop a contextually appropriate ethical framework regarding the use of AI in the practice of SLP in Pakistan. This consisted of a three-phase sequential study design: an online survey of 172 SLP practitioners, four focus groups with clinicians, academics, and students, and a two-round Delphi consensus process with 18 experts. The quantitative analysis consisted of reliability test, exploratory factor analysis, t-test, ANOVA and multiple linear regression. It was found that 55.2% of respondents were using large language models (LLDMs) in clinical documentation. Cultural-linguistic validity, data privacy and professional accountability were some of the larger ethical issues. The ethical concern scale was highly reliable ($\alpha = .91$) and factor analysis revealed four factors. The Delphi consensus led to the development of ten ethical guidelines, with at least 87% expert agreement. These guidelines offer a pragmatic approach for ethical application of AI in the context of SLP practice in Pakistan.



Introduction

AI (Artificial Intelligence) has become one of the most transformative technology advances in the 21st century healthcare system. In the field of speech-language pathology (SLP), AI has found its way into many aspects of clinical practice, such as automated speech and language assessment, dysarthria and dysphagia screening, augmentative and alternative communication (AAC), teletherapy delivery and production of clinical documentation. The potential to expand the availability of services to target populations that are underserved and the complex ethical issues raised by these technologies, including questions of accountability, consent, data governance, algorithmic bias, and access to services for all groups (Rao 2023; Millet et al. 2022) are all challenges facing the global SLP profession.

These challenges are particularly critical in the context of SLP profession, which is relatively new and still evolving in Pakistan. Pakistani SLP was founded in early 2000 and has no consolidated national regulatory body, has not adopted a code of ethics for its profession and lacks any published guidelines on the use of AI in clinical practice. While the Pakistan Association of Speech-Language Pathologists (PASLP) has taken steps to become professionalized, it has not yet developed any official position statements regarding the use of emerging technologies. In parallel, practices of SLP professionals in Pakistan are making use of AI tools, such as large language models, automated speech analysis platforms, and AI-assisted AAC devices, many of which have not been trained or normed for the local language or context in Pakistan, and are not fully aware of the ethical responsibilities.

SLP's context is particularly intricate when it comes to AI ethics in Pakistan. More than 70 languages and dialects are spoken in the country and Urdu is the national language but a big section of the populace speaks Punjabi, Sindhi, Pashto, Balochi and dozens of regional languages as mother tongues (Rehman 2011). This is because almost all AI speech and language tools that are commercially available have been created and tested in English or other high-resource languages, and their validity, reliability, and safety with Pakistani clients remains a fundamental question. In addition to linguistic differences, Pakistan also has a huge gap between urban and rural areas in terms of digital infrastructure, limited availability of resources in public health sites, and the sociocultural norms that exist in Pakistan regarding disability, privacy, and technology usage which can vary greatly from those in the West where the majority of AI ethics frameworks have been created.

At the global level several professional organizations have started to tackle AI ethics in SLP and related areas. American Speech-Language-Hearing Association (ASHA, 2023) has released guidelines for using AI in SLP practice and the World Health Organization (WHO, 2021) has issued guidelines for ethics in AI in health broadly. There is an international normative framework provided by the UNESCO (2021) Recommendation on the Ethics of Artificial Intelligence. These documents, however, have been written mostly about high-income, English-speaking environments and do not consider the particular context of SLP practice in Pakistan or other similar contexts. Therefore, an ethically grounded and empirically based framework of AI use in Pakistani SLP is needed and currently lacking.

This study aims to fill this gap by employing a three-phase mixed method design, including a survey, focus groups and Delphi consensus, and Delphi consensus to develop, validate and present ten guidelines for the ethical application of AI in the field of SLP in Pakistan. The study differs from studies based on the experiences of other professionals and centers its research on the experiences of Pakistani SLP professionals, academics and students and relies on internationally recognized ethical frameworks that have been adapted for the Pakistani context. It seeks to make a

contribution to Pakistani SLPs and to the broader conversation about the ethics of using AI in healthcare in a multilingual and culturally diverse context in low resource countries.

Aim and Research Questions

The overarching aim of this study was to develop contextually appropriate, empirically grounded ethical guidelines for the use of AI in SLP practice in Pakistan. The study was guided by the following research questions:

RQ1. What are the current patterns of AI tool use among registered SLP practitioners in Pakistan?

RQ2. What ethical concerns do Pakistani SLPs identify in relation to AI use in clinical practice?

RQ3. What themes emerge from qualitative exploration of AI ethics in Pakistani SLP practice?

RQ4. What ethical guidelines for AI use in Pakistani SLP achieve expert consensus through Delphi methodology?

RQ5. Which demographic and professional factors predict ethical AI practice readiness among Pakistani SLP practitioners?

Materials and Methods

Study Design

The design used was a three-phase sequential mixed methods from January to October 2025. Because it was a quantitative survey that first gave a broad overview of the use of and concerns about AI among SLPs, and focus groups were used to delve deeper into these concerns, the design worked well for this study. An expert consensus of contextually appropriate ethical guidelines was developed during the final Delphi phase. The three stages were an online cross sectional survey (Phase One), semi-structured focus groups (Phase Two), and a 2-round Delphi consensus process (Phase Three). An ethical clearance was granted by the Institutional Review Committee. Written informed consent was obtained from all participants and anonymity and data confidentiality were respected throughout the study.

Phase One: Online Survey

An online structured questionnaire was constructed and pretested with 10 additional SLP practitioners not part of the main sample. The questionnaire was adapted from the results of the pilot. The final questionnaire consisted of four sections: (a) demographic and professional background; (b) current use of AI in SLPs work functions; (c) ethical concerns with AI use (rated on a 5-point likert scale from 1 = not a concern to 5 = a very significant concern); and (d) open ended items about perceived barriers and facilitators of ethical use of AI. The questionnaire was given in the English language, which is used in the medical profession in Pakistan. It was disseminated electronically using PASLP networks, SLP university departments and professional groups on WhatsApp.

Those who had a recognised qualification in SLP and working or actively engaged in SLP related activity in Pakistan were eligible to participate. Purposive and snowball sampling method was employed in the recruitment of respondents. There were 172 complete usable responses to the questionnaire received and analyzed. Demographic and AI usage statistics were analyzed using descriptive statistics. Cronbach's α was used to check for internal consistency. To check the suitability of the ethical concern items for factor analysis, Kaiser-Meyer-Olkin test and Bartlett's test of sphericity were used. For the purpose of determining ethical concern domains exploratory factor analysis was carried out with principal axis factoring with varimax rotation. The differences

between the groups were analyzed using an independent samples t-test and one-way ANOVA for training status, workplace setting, and level of professional experience in ethical AI practice readiness. Multiple linear regression was used to determine the factors that predicted the readiness to practice ethical AI. Data were analyzed quantitatively with SPSS Version 27.

Phase Two: Focus Groups

Four semi-structured focus groups were held that included individuals who indicated interest in being involved in additional research during the survey. There were six to eight participants in each group. The groups were divided into: (a) clinical SLP practitioners (n = 7), (b) SLP academics and educators (n = 6), (c) early career SLPs (0–5 years) (n = 8), and (d) a mixed group (SLP students, carers of clients, allied health professionals) (n = 7). Different cities of Pakistan were able to participate in the focus groups conducted on Zoom. The sessions were of about 90 minutes duration each and were audio-taped upon request from the participants.

The recordings were verbatim transcribed and analysed using reflexive thematic analysis (RTA) proposed by Braun and Clarke (2006) with six stages. The transcripts were coded independently by two researchers who met to discuss coding differences. Any remaining differences were settled by discussion. All themes were examined in the entire data set for coherence, distinctiveness and in relation to participants' accounts. Two of the focus group participants were asked to do member checking to review the emerging theme summaries.

Delphi Consensus (Phase Three)

Participants were selected for the Delphi consensus process using an expert purposive sampling technique, which resulted in 18 participants. The panel included 6 senior clinical SLP practitioners who have experience with AI, 4 SLPs who are academics, 3 digital health / health technology specialists, 2 representatives from PASLP, 2 clinical ethicists and 1 AAC specialist. A consensus threshold of 75% was pre-specified, similar to the consensus threshold used in the traditional Delphi method (Hasson et al. 2000). The research team created 15 draft statements of ethical guidelines based on the results of Phase One and Phase Two, as well as international frameworks of AI ethics, such as WHO (2021), UNESCO (2021), ASHA (2023), and Beauchamp and Childress (2019).

During Round One, the panel members gave each statement a score from 1 = strongly disagree to 5 = strongly agree and included open-ended comments. Items with 75% or more agreement were kept as is; items below the 75% were changed and then re-presented in Round Two. Five new items were written in Round Two based on qualitative comments written in Round One. Ten statements after two rounds had obtained 87% or more consensus and were identified as the final ethical guidelines for the use of AI in Pakistani SLP practice.

Results

Phase One: Survey Findings

Table 1: Demographic and Professional Characteristics of Survey Respondents (N = 172)

Characteristic	Category	n (%)
Gender	Female	112 (65.1%)
	Male	55 (31.9%)
	Other/prefer not to say	5 (2.9%)
Professional Experience	< 2 years	38 (22.1%)
	2–5 years	64 (37.2%)

	6–10 years	45 (26.2%)
	> 10 years	25 (14.5%)
Work Setting	Hospital/clinical	78 (45.3%)
	Private practice	49 (28.5%)
	Academic/university	31 (18.0%)
	NGO/community	14 (8.1%)
City/Region	Karachi	52 (30.2%)
	Lahore	48 (27.9%)
	Islamabad/Rawalpindi	37 (21.5%)
	Peshawar/KPK	21 (12.2%)
	Other	14 (8.1%)

Note. Percentages may not sum to 100 due to rounding.

Table 1 provides the number of SLP practitioners who responded to the survey (172). The sample consisted mostly of females (65.1%) as is the case in the Pakistani SLP profession. Similarly, practitioners with a background of two to five years of experience accounted for the largest group of professional experience (37.2%) which is in keeping with the relative youth of the profession in Pakistan. The representation of the participants was from all four cities and regions with the highest representation being between Karachi with 30.2% and Lahore with 27.9%.

Table 2: Current AI Tool Usage Among Pakistani SLP Practitioners (N = 172)

AI Tool / Application	Currently Using (%)	Aware but Not Using (%)
AI-based speech assessment apps (e.g., voice analysis)	34.9%	41.3%
Teletherapy platforms with AI features	28.5%	38.4%
AAC with AI/ML components	22.1%	44.2%
AI-assisted language screening tools	19.8%	37.2%
Automated dysphagia risk assessment	11.6%	29.7%
AI-generated therapy plans/materials	41.3%	33.1%
Large language models (e.g., ChatGPT) for clinical documentation	55.2%	28.5%

Note. Percentages represent respondents selecting "currently using" or "aware but not currently using" from a three-option item (third option: "not aware of this tool type").

The findings from the survey showed that the use of AI tools is quite common but not uniform among all respondents (Table 2). The most widely used AI application was large language models, like ChatGPT, for clinical documentation (55.2%), followed by AI therapy plans or materials (41.3%) and AI-based speech assessment applications (34.9%). Knowledge of using AI tools was also generally more prevalent than usage across all categories, suggesting that many practitioners are either just beginning or are hesitant to engage with AI. Just 9.3% of the respondents said they had received any formal training during their SLP training and/or continuing education to learn how to use AI.

Table 3: Ethical Concerns Related to AI Use in SLP Practice: Rankings and Mean Scores (N = 172)

Rank	Ethical Concern	Mean Score (1–5)	% Rated High/Very High
1	Lack of cultural and linguistic validity of AI tools for Pakistani populations	4.71	93.0%
2	Data privacy and confidentiality of patient information	4.68	91.3%
3	Unclear accountability when AI produces erroneous clinical outputs	4.62	89.5%
4	Insufficient informed consent processes for AI-assisted intervention	4.55	87.2%
5	Risk of over-reliance on AI reducing clinical judgment	4.49	84.9%
6	Digital divide and inequitable access to AI tools	4.44	83.1%
7	Absence of national regulatory or professional guidelines on AI	4.41	82.6%
8	Algorithmic bias against minority language/dialect groups	4.38	81.4%
9	Insufficient training of SLPs in AI tool evaluation and use	4.29	79.1%
10	Ethical challenges in AI use with vulnerable populations (children, elderly)	4.21	76.7%

Note. Items rated on a 5-point Likert scale (1 = not a concern; 5 = very significant concern). Mean scores calculated from all valid responses.

Respondents rated ten ethical concern items (Table 3). The highest-rated concern was the lack of cultural and linguistic validity of AI tools for Pakistani populations ($M = 4.71$; 93.0% rated high or very high concern), followed by data privacy and confidentiality ($M = 4.68$; 91.3%) and unclear accountability for erroneous AI outputs ($M = 4.62$; 89.5%). All ten items received mean scores above 4.20, indicating that ethical concerns about AI were broadly and strongly held across the respondent group.

Table 4: Reliability Analysis of Study Scales (N = 172)

Scale	Number of Items	Cronbach α
Ethical Concerns Related to AI	10	.91
AI Tool Usage	7	.83
Ethical AI Practice Readiness	8	.89
Perceived Training Need	5	.86

Note. Cronbach α values above .70 were considered acceptable for research purposes (George and Mallery 2003).

Cronbach α coefficients indicated acceptable to excellent internal consistency across all key scales (Table 4). The ethical concern scale demonstrated excellent reliability ($\alpha = .91$), while the AI usage scale ($\alpha = .83$), ethical AI practice readiness scale ($\alpha = .89$), and perceived training need scale ($\alpha = .86$) also showed strong reliability. These results confirm that the survey items were internally consistent and suitable for further inferential analyses.

Table 5: KMO and Bartlett Test for Ethical Concern Items

Test	Value
Kaiser-Meyer-Olkin Measure of Sampling Adequacy	.876
Bartlett Test of Sphericity: Approx. chi-square	842.36
df	45
p	< .001

Note. A KMO value above .80 indicates meritorious sampling adequacy (Kaiser 1974).

Before conducting EFA, the KMO measure and Bartlett test of sphericity were examined (Table 5). The KMO value was .876, indicating meritorious sampling adequacy, and Bartlett's test was statistically significant ($\chi^2(45) = 842.36$, $p < .001$), confirming that the ethical concern items were sufficiently intercorrelated to proceed with factor analysis.

Table 6: Rotated Factor Structure of Ethical Concerns Related to AI Use

Factor	Main Items Loading on Factor	Eigenvalue	Variance Explained
Factor 1: Cultural and Linguistic Validity	Language validity, dialect bias, cultural relevance	3.18	31.8%
Factor 2: Privacy and Accountability	Data privacy, confidentiality, clinical responsibility	1.46	14.6%
Factor 3: Professional Competence	AI training, over-reliance, clinical judgement	1.21	12.1%
Factor 4: Equity and Access	Digital divide, vulnerable groups, rural access	.99	9.9%
Total Variance Explained			68.4%

Note. Principal axis factoring with varimax rotation. Only factor loadings $\geq .40$ were retained conceptually.

EFA using principal axis factoring with varimax rotation extracted four interpretable factors explaining 68.4% of total variance (Table 6). These were labelled: Factor 1 — Cultural and Linguistic Validity (eigenvalue = 3.18; 31.8% variance); Factor 2 — Privacy and Accountability (eigenvalue = 1.46; 14.6%); Factor 3 — Professional Competence (eigenvalue = 1.21; 12.1%); and Factor 4 — Equity and Access (eigenvalue = .99; 9.9%). This four-factor structure maps directly onto the core ethical domains identified in the qualitative and Delphi phases of the study.

Table 7: Differences in Ethical AI Practice Readiness

Comparison	Group	n	Mean	SD	Test Statistic	p
AI Training	Formal AI training received	16	4.18	.46	t = 4.27	< .001
	No formal AI training	156	3.42	.67		
Work Setting	Hospital/clinical	78	3.52	.64	F = 5.14	.002
	Private practice	49	3.79	.59		
	Academic/university	31	3.88	.53		
	NGO/community	14	3.31	.71		
Experience	< 2 years	38	3.28	.69	F = 4.62	.004
	2–5 years	64	3.49	.62		

	6–10 years	45	3.83	.57		
	> 10 years	25	3.71	.60		

Note. *Ethical AI practice readiness was measured on a 5-point scale (1 = very low readiness; 5 = very high readiness). Higher scores indicate stronger readiness for ethical AI use in practice.*

Table 7 shows the group differences in the participants' readiness to practice ethical AI as clinicians. An independent samples t test showed that SLPs who participated in formal AI training ($n = 16$) scored higher on the ethical AI practice readiness scores ($M = 4.18$, $SD = .46$) than SLPs who had not received formal training ($n = 156$; $M = 3.42$, $SD = .67$), $t(170) = 4.27$, $p < .001$. This indicates that those who have received formal instruction in AI-related training are more prepared to use AI ethically in their clinical practice.

Furthermore, a one-way ANOVA showed that there were statistically significant differences between work settings in terms of ethical AI practice readiness: $F(3, 168) = 5.14$, $p = .002$. The highest readiness scores were reported by practitioners in academic or university settings ($M = 3.88$), suggesting that the institutional exposure to research, teaching, and emerging technologies can help foster further ethical preparedness related to AI. There were also differences in readiness by professional experience groups, $F(3, 168) = 4.62$, $p = .004$, as practitioners with 6–10 years of experience indicated the greatest readiness ($M = 3.83$). The results suggest that formal education, work environment, and job experience can be significant factors in the development of ethical preparedness in the use of AI in SLP.

Table 8: Multiple Linear Regression Predicting Ethical AI Practice Readiness ($N = 172$)

Predictor	B	SE B	β	t	95% CI	p
Constant	2.14	.31	—	6.90	[1.53, 2.75]	< .001
Formal AI training	.54	.12	.36	4.50	[.30, .78]	< .001
AI usage frequency	.29	.08	.28	3.63	[.13, .45]	< .001
Professional experience	.11	.05	.17	2.20	[.01, .21]	.029
Academic/university setting	.22	.10	.16	2.20	[.02, .42]	.030
Ethical concern score	.18	.07	.19	2.57	[.04, .32]	.011

Note. $R = .62$, $R^2 = .38$, adjusted $R^2 = .36$, $F(5, 166) = 20.36$, $p < .001$. The dependent variable was ethical AI practice readiness. 95% CI = 95% confidence interval for unstandardised coefficient B.

Table 8 presents the results of the multiple linear regression analysis conducted to identify predictors of ethical AI practice readiness among SLP practitioners. The overall regression model was statistically significant, $F(5, 166) = 20.36$, $p < .001$, with $R = .62$ and $R^2 = .38$. This means that the five predictors together accounted for 38% of the variance of the ethical AI practice readiness scores. Formal AI training was the most significant positive predictor of readiness for ethical AI use ($B = .54$, $\beta = .36$, $p < .001$), meaning that those practitioners who reported having received formal AI-related training were more likely to report a higher readiness for ethical AI use. The frequency of AI use was also a significant positive predictor ($B = .29$, $\beta = .28$, $p < .001$), indicating that more practical experience with AI tools was positively related to ethical preparedness.

In addition, ethical concern score was significantly predictive of readiness ($B = .18$, $\beta = .19$, $p = .011$), with practitioners who were more aware of the ethical concerns showing higher readiness to practice ethically with AI. As with the other variables, professional experience was a significant predictor, ($B = .11$, $\beta = .17$, $p = .029$), with higher levels of clinical experience positively associated with ethical readiness. Also, employment in an academic or university work environment accounted for a significant amount of variance in readiness scores ($B = .22$, $\beta = .16$, $p = .030$). The results indicated that formal AI training, frequent use of AI, ethical awareness,

professional experiences, and academic exposure are significant factors that influence the readiness of Pakistani SLPs in practice with ethical considerations in the use of AI.

Phase Two: Focus Group Findings

The focus group discussions were analyzed using a thematic approach, resulting in the identification of five overall themes that represent the views of the participants regarding the ethical use of AI as a tool in the work of speech-language pathologists. Each of these themes offers a detailed overview of the opportunities and challenges for AI adoption, and builds on the quantitative survey results. Examples of representative quotations are reported to exemplify each theme and to increase the trustworthiness and credibility of the qualitative analysis.

Theme 1: Linguistic Unsuitability of Existing AI Tools

The first major theme was the concern that the majority of the existing speech and language AI tools are not linguistically appropriate for speaking, writing, and reading native speakers of the Urdu language, who are often the focus of Pakistani clients. The participants were aware that these tools are predominantly created and tested in English-speaking and Western populations and thus, their direct use in Pakistan is restricted. Clinical practitioners described that AI tools are not effective for Urdu phonology, regional accents, and dialects.

This one senior clinician said, *“I did try an AI speech assessment tool with a kid from a Pashtun family, and the results were totally meaningless because the tool was not able to pick up his phonological system”* (P3, Clinical SLP). Likewise, another participant commented that such tools could wrongly interpret dialectal features as errors in speech, which could lead to missed assessment and planning for inappropriate intervention (P6, Clinical SLP). Also noted by the academic participants were the lack of locally validated AI speech and language tools. They also said that AI tools should not be used as medical decision making tool without testing on the samples of Pakistani language and culturally relevant communication pattern. An academic participant said, *“If you are talking about the output of AI in Pakistani languages, then without the validation of these languages, it will not be clinically reliable”* (P10, SLP Academic). Overall, participants considered linguistic inadequacy as a major ethical issue due to the impact of linguistic inadequacy on the accuracy of the assessment, clinical judgment and client safety.

Theme 2: Absence of Regulatory and Professional Guidelines

The second theme was related to the absence of any national or professional guidelines regarding the application of AI in the field of SLP in Pakistan. The focus groups participants were unsure of the ethical, legal, and professional limits of the use of AI in the assessment, intervention, documentation, and counselling process. One early career SLP explained, *“I was wanting to use an AI for an autism child and didn't know if it was ethical or if it needed special consent”* (P15, Early-Career SLP). One further participant commented that *“when AI gives a wrong result, there is nothing that is written in Pakistan on what to do with that”* (P18, Early-Career SLP). Participants expressed the sense that there is a lack of direction and that decisions made by doctors are left to the individual clinician and should be supported by professional standards and institutional policies. Academic participants also reported that the training programmes for SLPs offer little training on AI literacy, digital ethics, and technology evaluation. One teacher commented that *“students are playing with AI tools, but not part of a formal training process in judging the accuracy, safety, or ethical risks of the tools”* (P9, SLP Academic). The theme suggests that there is a requirement for formal ethical guidelines, curriculum development, and professional regulation for the use of AI in the practice of SLPs in Pakistan.

Theme 3: Consent, Transparency, and Informed Decision-Making

The third theme centered on informed consent and transparency issues in SLP practice with the aid of AI. Participants shared that there is a risk that many of the clients and caregivers may not fully appreciate how clients and caregivers can collect, analyses, store, and interpret clinical data with AI. For many of my clients, coming from a low income background, *if I explain that AI is analyzing my child's speech, they don't really understand what I'm saying*” (P5, Clinical SLP). Another participant commented that sometimes families do not ask questions when a clinician decides on a course of action; this can make informed consent challenging (P7, Clinical SLP). Additionally, participants raised the issue that the algorithms, training data, accuracy rates and data storage practices are not clearly specified for many commercial AI platforms. This opacity hinders SLPs' ability to communicate AI-driven suggestions to clients and to effectively review clinical outcomes. *“Without understanding how the tool achieved results, I will not be able to confidently communicate the results to parents”* (P12 SLP Academic). The participants thus emphasized that the use of AI in SLP should be accompanied by explanations, culturally tailored consent mechanisms, and transparent documentation of the role of AI in clinical decision-making.

Theme 4: Risk of Professional Deskilling and Over-Reliance

The fourth theme emphasized participants' worry about the use of AI tools becoming so pervasive that it undermines clinical reasoning and clinical judgment. Early-career SLPs and academic participants indicated that reports generated by AI, therapy goals, and therapy plans may be helpful, but can also cause dependency. *“If students are only given reports and goals created by AI, they might never be able to think clinically on their own”* (P11, SLP Academic). A second participant commented that there is no replacement for clinical observation, listening and the ability to develop a hypothesis (P14, Early-Career SLP). The importance of using AI to support the role of the qualified SLP was consistently emphasized. They expressed worries about a dependence on AI, which could lead to less confidence in the profession's practice, less diagnostic reasoning, and less responsibility for error if it happens. One clinician stated, *“Who fails if the AI recommends a different therapy goal and the child fails to improve?”* (P4, Clinical SLP). The theme reveals that the use of AI must always have a human clinical point of decision, with an increased focus on AI literacy, and human responsibility to clinical decisions.

Theme 5: Digital Inequity and Access

The fifth theme emphasized the potential for exacerbating the existing inequities of access to SLP services with the implementation of AI. The participants reported that there is a lack of infrastructure for the use of artificial intelligence in public sector and in rural clinical setting, such as internet connectivity, adequate devices, paid software and technical assistance. A clinician with experience in a resource limited setting commented, *“In many public hospitals, even the simplest digital facilities are limited, and so are the advanced AI tools – they are not practically available”* (P2, Clinical SLP). Another participant commented that there are many families and institutions for whom the cost of AI-based subscription programs is too much (P21, Early-Career SLP). Participants felt that the use of AI tools to enhance access to SLP services will only be possible if the tools are affordable, validated, and offered outside of private urban clinics. They said if AI is restricted to *“wealthier clients”* it could deepen the urban/rural divide and the public/private divide. *“Technology should help lessen inequity, not be another disincentive for poor families”* (Allied Health Professional, P26). Overall, participants highlighted the need for equity, affordability, access for the public sector, and safeguarding vulnerable groups to ensure that AI is implemented ethically in Pakistan.

Phase Three: Delphi Consensus Guidelines

Table 9: Ethical Guidelines for AI Use in Pakistani SLP Practice: Delphi Consensus Results

#	Ethical Principle	Guideline Statement	Consensus (%)
1	Cultural and Linguistic Validity	SLPs must evaluate AI tools for validity and appropriateness with Pakistani language populations (Urdu, Punjabi, Sindhi, Pashto, etc.) before clinical use.	97%
2	Informed Consent	Clients and caregivers must receive clear, culturally appropriate explanations of AI tool use, including its limitations, before consent is obtained.	96%
3	Data Privacy and Security	Patient data processed by AI systems must comply with applicable Pakistani data protection law and international best practice (e.g., GDPR principles).	95%
4	Professional Accountability	Clinical responsibility for AI-assisted decisions remains with the qualified SLP at all times; AI outputs must not substitute professional clinical judgment.	95%
5	Transparency	SLPs must be able to explain in accessible terms how an AI tool reached its output, and must document this in clinical records.	93%
6	Equitable Access	AI tools should not deepen existing disparities; SLPs and institutions should advocate for equitable access to validated AI tools across urban and rural settings.	91%
7	Competence and Training	SLPs must only use AI tools for which they have received adequate training; AI literacy should be incorporated into undergraduate and continuing professional development curricula.	90%
8	Non-Maleficence	SLPs must critically evaluate AI tools for potential harms including algorithmic bias, diagnostic error, and inappropriate recommendations before clinical deployment.	89%
9	Beneficence	AI tools should be adopted only when there is reasonable evidence that they improve clinical outcomes, efficiency, or access to services for Pakistani clients.	88%
10	Child and Vulnerable Population Protection	Additional safeguards must be applied when using AI tools with children, elderly individuals, or those with cognitive or communication impairments.	87%

Note. Consensus defined as $\geq 75\%$ of panellists rating the statement as 4 (agree) or 5 (strongly agree). Final percentages reflect Round Two results. All retained items exceeded the pre-specified 75% threshold; the lowest final consensus was 87%.

Table 9 presents the final ethical guidelines for the use of AI in Pakistani SLP practice, developed through a two-round Delphi consensus process. The ten guideline statements achieved high level

of expert agreement ranging from 87% to 97% with all statements being above the threshold of 75% agreement. The highest consensus was reported for cultural and linguistic validity (97%), meaning that experts thought that cultural appropriateness and the use of local language were the most critical factors before using AI tools in clinical settings. The experts also shared high consensus on informed consent, data privacy, professional accountability and transparency, indicating that these are important ethical concerns for clients, privacy, and the responsibility of the clinician. Also with high consensus were guidelines in the areas of equity of access, competence and training, non-maleficence, beneficence, and the protection of vulnerable populations. The overall results of the Delphi process suggested that the ethical use of AI in Pakistani SLP practice should be guided by the following four major realms: Validity and safety, Rights and consent, Professional responsibility, and Equity of access.

Discussion

This study, to the researchers' best knowledge, is one of the first empirical studies in the region of SLP practice within Pakistan that provides descriptive evidence of current use of AI, ethical concerns, and a consensus-based set of ethical guidelines validated by experts. The results paint a picture of a profession on the precipice of change: AI use is already pervasive, with significant issues of substance on ethics broadly accepted, but professional “infrastructure” to regulate AI use is almost non-existent.

The Cronbach α values were high, indicating the reliability of the survey scales, and the EFA results indicated that the ethical concerns of the Pakistani SLPs, and not just one vague general concern, are grouped into distinct four domains: cultural-linguistic validity, privacy/accountability, professional competence, and equity/access. It is important to consider how an ethical guideline and training programme can be designed, as only one or two of these domains will not cover the full spectrum of practitioners' concerns. The regression results also provide a greater clarificatory understanding of the formal AI training and actual AI exposure as key modifiable predictor of ethical readiness, which further argues that ethical AI usage in Pakistan cannot be limited to personal vigilance and caution, but requires a structured approach to AI training, institutional policy, and professional standards.

The finding that over half of respondents (55.2%) are using large language models for clinical documentation is particularly significant and is This finding suggests substantial informal adoption of general-purpose AI tools (Rao 2023; Millet et al. 2022). This probably comes down to the fact that formal tools made to fit the SLP use case are costly or unavailable, and free general purpose AI tools are a convenient alternative. There are significant risks with this pattern: outputs may be clinically inaccurate, and such models are neither validated for clinical documentation accuracy in SLP nor subject to healthcare data protection requirements. This disparity between the high prevalence of informal AI use and the lack of near-complete professional learning opportunities of AI (9.3% reported professional development around AI) is a clear priority for professional learning.

The top-ranked ethical issue, which was lack of cultural and linguistic validity of algorithm for Pakistani populations, is directly related to the literature on algorithmic bias and health equity in international literature. But Obermeyer and her colleagues (2019) showed that commercial health AI algorithms, trained on non-representative data, incorrectly underestimate health needs in minority populations, and similar risks exist for Pakistani language communities whose usage patterns and norms of communication are not included in any existing commercial language AI speech tool. This risk is addressed directly by the Delphi guideline which mandates that before using it clinically it must be validated for Pakistani language populations.

The consent and transparency issues that emerged from the survey and focus group align with this fundamental tension between AI-assisted healthcare: meaningful informed consent demands AI literacy which many clients and practitioners do not possess, and transparency requirements with respect to proprietary AI systems are challenging. This is called "ethics of algorithms" problem, according to Mittelstadt et al. (2016). This challenge is worsened in the Pakistani context, as average digital literacy among clients is low, and there are cultural aspects of the clinician-client relationship that makes consent processes difficult, which is why culturally adapted consent processes in Delphi Guideline 2 need to be outlined.

The concept of professional deskilling echoes the discussions surrounding AI and professional skills. While Topol (2019) contended that AI in healthcare could "liberate" clinicians from mundane tasks to concentrate on clinically complex reasoning, critics highlighted the need to make strategic investments in the pedagogical process, especially for teaching basic skills before introducing clinicians to AI support (Wachter and Shah 2021). This worry by Pakistani SLP practitioners about the possible negative impact of relying on AI on the development of clinical reasoning skills is directly linked to curriculum design and directly covered by Delphi Guideline.

The digital equity theme reflects the broader challenge that AI tools created in high-income environments may create a new inequity if not designed explicitly with equity in mind for use in resource-limited environments (Char et al. 2018). In Pakistan, the SLP workforce is focused in urban areas, and clients in the public sector often have limited digital access, making equity conscious deployment imperative and an ethical duty (as stated in Delphi Guideline 6).

Some of the limitations of this study should be noted. The practitioners surveyed may be over-represented in the sample because it was drawn from the professional network and may include professionals who are already interested in technology or ethics. The sample for the focus group was predominantly from an urban area. The Delphi panel was capped at 18 members and the views of client groups and policy makers outside of SLP were not formally represented. It is recommended that future studies should involve larger more representative samples, involve client/career voices, and study implementation and impact of guidelines in clinical practice.

Conclusion

This study shows that the use of AI in the SLPs' practice is already substantial, with little ethical regulation. SLP practitioners in Pakistan are deeply concerned about the ethical implications of using AI, including its impact on linguistic validity, data privacy, professional responsibility, consent, and accessibility. A rigorous mixed-methods approach (survey, focus groups and Delphi consensus) has been used to produce ten contextually relevant ethical guidelines, which have been validated by expert consensus. These guidelines offer a practical template for PASLP and other professional organizations to create formal position statements, as well as for SLP training programs to integrate AI ethics into their curricula, and for clinical institutions to develop AI governance policies. In Pakistan, the landscape is marked by linguistic diversity, resource limitations, digital disparities, and the growth of AI facilities and infrastructure among professionals, making it a representative example of a vast region where AI is progressing at a rate outpacing ethical frameworks and governance. The guidelines and methodology framed here thus provide a replicable model for other SLP associations and health professions in South Asia and similar countries across the world. The ethical use of AI in Pakistani SLP is not about eliminating the technology, but rather about regulating, informing and making it equitable to use, and to do so with the guidance of the users, the clients and the communities it's meant to serve.

Recommendations

1. **Professional bodies**, particularly PASLP, should develop and adopt national ethical guidelines for AI use in speech-language pathology practice.
2. **Universities** should integrate AI ethics, digital health, and responsible AI use into undergraduate and postgraduate SLP curricula.
3. **Healthcare institutions** should establish policies governing AI-assisted assessment, documentation, informed consent, and data protection.
4. **AI developers** should validate speech and language technologies using Pakistani languages and culturally diverse datasets before clinical implementation.
5. **Continuing professional development (CPD)** programmes should provide regular AI literacy and ethics training for practicing SLPs.
6. **Future research** should evaluate the implementation, feasibility, and clinical effectiveness of the proposed ethical guidelines across different healthcare settings in Pakistan

References

1. American Speech-Language-Hearing Association (ASHA). 2023. Artificial Intelligence: Considerations for Speech-Language Pathology Practice. Rockville, MD: ASHA. <https://www.asha.org/practice/artificial-intelligence/>
2. Barocas, S., M. Hardt, and A. Moritz. 2019. Fairness and Machine Learning: Limitations and Opportunities. Cambridge, MA: MIT Press.
3. Beauchamp, T. L., and J. F. Childress. 2019. Principles of Biomedical Ethics. 8th ed. Oxford: Oxford University Press.
4. Braun, V., and V. Clarke. 2006. "Using Thematic Analysis in Psychology." *Qualitative Research in Psychology* 3 (2): 77–101. <https://doi.org/10.1191/1478088706qp063oa>
5. Char, D. S., N. H. Shah, and D. Magnus. 2018. "Implementing Machine Learning in Health Care—Addressing Ethical Challenges." *New England Journal of Medicine* 378 (11): 981–983. <https://doi.org/10.1056/NEJMp1714229>
6. Creswell, J. W., and V. L. Plano Clark. 2017. Designing and Conducting Mixed Methods Research. 3rd ed. Thousand Oaks, CA: SAGE Publications.
7. George, D., and P. Mallery. 2003. SPSS for Windows Step by Step: A Simple Guide and Reference. 4th ed. Boston, MA: Allyn and Bacon.
8. Hasson, F., S. Keeney, and H. McKenna. 2000. "Research Guidelines for the Delphi Survey Technique." *Journal of Advanced Nursing* 32 (4): 1008–1015. <https://doi.org/10.1046/j.1365-2648.2000.t01-1-01567.x>
9. Kaiser, H. F. 1974. "An Index of Factorial Simplicity." *Psychometrika* 39 (1): 31–36. <https://doi.org/10.1007/BF02291575>
10. Millet, B., J. Boster, and V. Ramanarayanan. 2022. "Artificial Intelligence Applications in Speech-Language Pathology: A Systematic Review." *American Journal of Speech-Language Pathology* 31 (5): 2053–2071. https://doi.org/10.1044/2022_AJSLP-21-00348
11. Mittelstadt, B. D., P. Allo, M. Taddeo, S. Wachter, and L. Floridi. 2016. "The Ethics of Algorithms: Mapping the Debate." *Big Data and Society* 3 (2): 1–21. <https://doi.org/10.1177/2053951716679679>
12. Obermeyer, Z., B. Powers, C. Vogeli, and S. Mullainathan. 2019. "Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations." *Science* 366 (6464): 447–453. <https://doi.org/10.1126/science.aax2342>

13. Rao, P. K. 2023. "Artificial Intelligence and the Future of Speech-Language Pathology." Perspectives of the ASHA Special Interest Groups 8 (1): 195–205. https://doi.org/10.1044/2022_PERSP-22-00064
14. Rehman, T. 2011. Punjabi Language and Pakistani Nationalism. Islamabad: National Language Authority.
15. Topol, E. J. 2019. "High-Performance Medicine: The Convergence of Human and Artificial Intelligence." Nature Medicine 25 (1): 44–56. <https://doi.org/10.1038/s41591-018-0300-7>
16. UNESCO. 2021. Recommendation on the Ethics of Artificial Intelligence. Paris: UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000381137>
17. Wachter, R. M., and N. H. Shah. 2021. "Suffering in Silence: How Electronic Health Records Are Hurting Us." New England Journal of Medicine 381 (12): 1174–1179. <https://doi.org/10.1056/NEJMms2009175>
18. World Health Organization (WHO). 2021. Ethics and Governance of Artificial Intelligence for Health. Geneva: WHO. <https://www.who.int/publications/i/item/9789240029200>