Supplemental Information

Preliminary Interviews

As a purely exploratory investigation to set the groundwork for developing AI tools, eight 1:1 in-depth interviews were conducted with both Spring Health and non Spring Health therapists to understand their opinion of AI tools prior to implementation of session recording and AI-notetaking. Interviews focused on trust, education, and use cases hypothesized to contribute to utilization. These interviews were not subject to formal qualitative analysis; rather they served to highlight only general opinion, trends, and AI tool experience among providers.

Some providers had used AI tools such as chatGPT in their personal life and watched the space of AI closely. They were intrigued by the progression of the technology and were eager to test out AI tools professionally to get a sense of their effectiveness and usefulness. This suggests there is a segment of therapists who are more intrinsically motivated to try out new AI technologies and who are more comfortable using such tools. (Although we did not collect user experience data from patients, the same is likely true for a segment of patients).

However, there was also a segment of providers who had little experience with AI and were much more reluctant to adopt such technologies. They expressed doubt that session recordings could ever be truly temporary or deleted, believing that “everything leaves a footprint.” They worried about breaches, misuse of PHI, or lack of transparency about where data is stored, for how long, and who has access.

Overall, regardless of their initial attitude about AI, providers feared that recording sessions could undermine the trust and openness that therapy relies on. They worried that in-session AI tools or knowledge that a session was being recorded might distract therapists, making them less present, or even harm client perceptions of their therapist's professionalism.

Supplemental Data

Supplemental Table 1. Time spent on clinical note as a function of AI-notetaking usage

	Time (mins) spent on clinical note		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	27.30	26.93 – 27.66	**<0.001**
AI tool usage [session recorded, but AI-notes not copied]	-2.69	-2.96 – -2.42	**<0.001**
AI tool usage [session recorded and AI-notes copied]	-8.54	-8.73 – -8.36	**<0.001**
Random Effects			
σ^2	761.82		
$\tau_{00\ PROVIDER\_ID}$	248.53		
ICC	0.25		
$N_{\ PROVIDER\_ID}$	8619		
Observations	752479		
Marginal R^2 / Conditional R^2	0.014 / 0.256		

Supplemental Table 2. Time to submit and sign and submit clinical notes as a function of AI-notetaking usage.

	Hours to submit note		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	42.22	41.34 – 43.09	**<0.001**
AI tool usage [session recorded, but AI-notes not copied]	-4.54	-5.13 – -3.96	**<0.001**
AI tool usage [session recorded and AI-notes copied]	-8.18	-8.80 – -7.56	**<0.001**
Random Effects			
σ^2	1184.86		
$\tau_{00\ PROVIDER\_ID}$	1567.10		
$\tau_{11\ PROVIDER\_ID.AI\ tool\ usage\ [session\ recorded,\ but\ AI\text{-}notes\ not\ copied]}$	142.49		
$\tau_{11\ PROVIDER\_ID.AI\ tool\ usage\ [session\ recorded\ and\ AI\text{-}notes\ copied]}$	311.83		
ρ_{01}	-0.31		
	-0.50		
ICC	0.55		
$N_{PROVIDER\_ID}$	8532		
Observations	689132		
Marginal R^2 / Conditional R^2	0.005 / 0.551		

Supplemental Table 3. Clinical notes submission “after hours” as a function of AI-notetaking usage.

	After hours submission		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	0.081	0.077 – 0.085	**<0.001**
AI tool usage [session recorded, but AI-notes not copied]	0.888	0.861 – 0.915	**<0.001**
AI tool usage [session recorded and AI-notes copied]	0.813	0.796 – 0.830	**<0.001**
Random Effects			
σ^2	3.29		
$\tau_{00\ \text{PROVIDER\_ID}}$	3.74		
ICC	0.53		
$N_{\text{PROVIDER\_ID}}$	8602		
Observations	736780		
Marginal R^2 / Conditional R^2	0.001 / 0.533		

Supplemental Table 4. Clinical Symptom Improvement

	Depression (PHQ 9)			Anxiety (GAD 7)		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	15.38	15.28 – 15.49	**<0.001**	14.58	14.49 – 14.66	**<0.001**
AI tool usage [used AI]	-0.07	-0.22 – 0.08	0.358	-0.12	-0.24 – 0.00	0.058
time in treatment (log-days)	-1.47	-1.50 – -1.44	**<0.001**	-1.34	-1.37 – -1.31	**<0.001**
AI tool usage [used AI] * time in treatment (log-days)	-0.03	-0.07 – 0.02	0.265	-0.01	-0.05 – 0.04	0.750
Random Effects						
σ^2	11.79			10.83		
τ_{00}	8.43 $_{member\_id}$			2.39 $_{member\_id}$		
	0.42 $_{provider\_id}$			0.12 $_{provider\_id}$		
τ_{11}	0.70 $_{member\_id:time\_in\_treatment}$			0.49 $_{member\_id.time\_in\_treatment}$		
ρ_{01}	0.03 $_{member\_id}$			0.87 $_{member\_id}$		
ICC	0.59			0.54		
N	14247 $_{member\_id}$			14104 $_{member\_id}$		
	4212 $_{provider\_id}$			4202 $_{provider\_id}$		
Observations	43358			42690		
Marginal R^2 / Conditional R^2	0.226 / 0.682			0.225 / 0.646		