



Innovating Responsibly
For Mental Healthcare

WHITE PAPER

The bottleneck is use, not collection

*Making longitudinal client data actionable in
psychotherapy*

Cognitai Research Department

Cognitai Health Inc., Montréal, Québec, Canada

Version 1.0, October 2026

Evidence review since August 2025

Research conducted: August 2025 to October 2026

Summary

A common view holds that clients will not provide data between psychotherapy sessions. This narrative review of peer-reviewed sources from 2020 onward finds that the view comes largely from self-guided apps. In clinical care, clients provide repeated data, and feeding it back to clinicians is associated with small average benefits and larger ones for clients who are not on track. The bottleneck is use: turning that data into something a therapist can interpret, trace to its source and act on. The paper proposes seven design principles, offered as hypotheses, and states the limits of the evidence.

Key findings

- 01** Clients in clinical care provide repeated data. In a randomized trial of adults with major depressive disorder, participants completed a median of 9 of 12 weekly questionnaires, and wearable data met the study's threshold on a mean of 74 of 84 days.
White et al., 2024
- 02** How measurement is built into care, and whether a person supports it, shapes completion. In one outpatient program, PHQ-9 completion rose from 5% to 66% of visits in the first rollout phase and to 60% in the second. Human guidance raised completion in a meta-analysis of 22 studies ($g = 0.29$).
Steidtmann et al., 2025; Musiat et al., 2022
- 03** Feeding data back to clinicians is associated with small average benefits ($d = 0.14$ to 0.15) and larger ones for clients who are not on track ($d = 0.29$ in one meta-analysis).
Rognstad et al., 2023; de Jong et al., 2021
- 04** Clinicians value measurement but use it inconsistently. In a Canadian survey, 83.3% of 108 clinicians agreed that measurement-based care is important for clinical judgement, yet 39.8% used it routinely. One large implementation study warns that gathering data without integrating it into care "may even be harmful."
Cheung et al., 2023; Forand et al., 2025
- 05** Seven design principles follow, offered as hypotheses and not as findings: clinician control, interpretable summaries, traceability to the source, association and not causation, minimal client burden, privacy and client choice, and measuring usefulness. The evidence base is heterogeneous, partly observational and partly industry-authored.
Joyce et al., 2023; Barkham et al., 2023

How to cite (APA 7)

Cognitai Research Department. (2026). The bottleneck is use, not collection: Making longitudinal client data actionable in psychotherapy [White paper]. Cognitai Health Inc.
<https://cognitai.org/research/bottleneck-use-not-collection>

About this paper

The Cognitai Research Department began this evidence review in August 2025; this paper is its first public edition, updated with studies published through 2025. It is a narrative review, not a systematic review.

Disclosure

This paper was written by the Cognitai Research Department, the research function of Cognitai Health Inc., a company developing software in the area it discusses (clinician-facing tools for organizing information gathered between psychotherapy sessions). This is a potential conflict of interest. No external funding was received for this work. The paper reports no new data, and no performance claims are made for any product.

About the Cognitai Research Department

The Cognitai Research Department is the research function of Cognitai Health Inc., in Montréal. It studies how longitudinal client information, gathered within and between sessions, can become interpretable, traceable and clinically useful. Its focus is helping mental healthcare become more evidence-informed and data-driven, with one goal: **maximizing continuity, precision, and personalization in mental healthcare while minimizing clinician workload**. Its areas of interest include measurement-based care, the therapeutic alliance, between-session continuity, documentation burden, and the responsible use of AI in clinical practice. The department draws on the peer-reviewed literature, structured discovery with clinicians and clients, and the evaluation of pilots in real clinical settings. Its work rests on principles shared with Cognitai's approach to care: the clinician remains in control, every insight is traceable to its source, associations are never presented as causes, and privacy is protected by design. It publishes white papers and evidence briefs that state the limits of the evidence, and it discloses that Cognitai develops clinician-facing software in this area. It receives no external funding.

Abstract

A common view holds that clients will not provide data between psychotherapy sessions, a view drawn largely from self-guided apps. We argue that this view transfers poorly to clinical care and that the more consequential bottleneck lies downstream: turning longitudinal client data into information that therapists can interpret and act on. In this narrative review of peer-reviewed sources from 2020 onward, we first summarize evidence that clients do provide repeated self-report and passive data when measurement is embedded in care or supported by people, including a randomized trial in which participants completed a median of 9 of 12 weekly questionnaires, an outpatient program in which questionnaire completion rose from 5% to 60% or 66% of visits across two rollout phases, and a meta-analysis showing that human guidance increases completion. Second, we review meta-analytic and pragmatic evidence that feeding these data back to clinicians is associated with small overall benefits and larger benefits for clients who are not on track. Third, we describe the usability gap: clinicians value measurement but use it inconsistently, cite efficiency and ease of use as barriers, and carry a heavy documentation burden; one large implementation study warns that collecting data without integrating it into care “may even be harmful.” We then propose seven design principles for making longitudinal data clinically useful, including clinician control, interpretable summaries, traceability of every insight to its source, and the explicit framing of associations as associations. We close with the limits of a heterogeneous, partly observational, partly industry-authored evidence base.

Keywords: measurement-based care, routine outcome monitoring, measurement feedback systems, digital mental health, clinician burden, explainable artificial intelligence

1 Introduction

Psychotherapy is organized around a weekly or less frequent hour, yet most of what therapy is meant to change happens outside it. Sleep, mood, conflict, avoidance, and the practice of new skills unfold across the remaining days of the week, and therapists usually learn about them through the client’s recall at the start of the next session. Measurement-based care (MBC) and routine outcome monitoring (ROM) aim to bring some of that time into view by collecting standardized self-reports at regular intervals and feeding the results back to clinicians (Barkham et al., 2023). Digital tools now allow the same logic to extend further, to short check-ins between sessions, digital exercises, and, where clients choose to share them, passive data from wearable devices.

A frequent objection to such approaches is that clients simply will not provide the data. The objection is usually supported by evidence from self-guided consumer apps and from trials of standalone digital interventions. A review of engagement with digital mental health interventions concluded that “patient engagement with DMHIs in real-world settings is often insufficient to see clinical benefit” (Lipschitz et al., 2023, p. 119). A meta-analysis of 70 randomized controlled trials (RCTs) of smartphone-delivered interventions found mean study attrition of 24.1% (95% CI [19.3, 29.6]) at short-term follow-up and 35.5% (95% CI [26.7, 45.3]) at longer-term follow-up, with intervention usage that “consistently declined over the course of the trial” (Linardon & Fuller-Tyszkiewicz, 2020, p. 1).

These findings are important, but they describe a different situation from the one most clients in psychotherapy are in. A person who downloads an app alone has no therapist expecting the data, no session in which it will be discussed, and no relationship in which providing it has a purpose. Clients in clinical care are in the opposite position. Our central argument is that, in that context, collection is no longer the main constraint. The more consequential and less studied problem is what happens after collection: whether therapists can read, interpret, trust, and act on longitudinal data within the time and cognitive budget of real practice.

This article is a narrative review and perspective. It is not a systematic review. We restricted the evidence to peer-reviewed sources published from 2020 onward so that the argument rests on the current generation of digital and implementation research, and we report each statistic as the original authors defined it. Throughout, we use *client* for the person receiving care, including when describing studies whose authors used *patient*; quoted text is left unchanged. The article proceeds in four steps: evidence that clients provide data in clinical contexts; evidence that feedback of those data to clinicians improves outcomes, particularly for clients who are not on track; evidence of a usability gap between collection and clinical use; and a set of design principles for closing that gap.

2 Clients provide data in clinical contexts

Repeated self-report and passive data over weeks

The most direct test of whether people with depression will keep providing data over time comes from remote measurement research. White et al. (2024) conducted a two-arm RCT with 100 adults with major depressive disorder enrolled in the RADAR-MDD remote measurement program. Over 12 weeks, participants were asked to complete the eight-item Patient Health Questionnaire (PHQ-8) weekly and to wear a Fitbit device. The primary outcome, the percentage of the 12

weekly PHQ-8 questionnaires completed, was 69% in the control arm and 68% in the intervention arm ($P = .83$); the intervention tested to increase engagement did not change completion. Participants completed a median of 9 of 12 questionnaires (interquartile range 6 to 10), although only 7% completed all 12. Wearable data meeting the study's threshold of at least two recordings per day were available on a mean of 74 of 84 days.

Two features of this study matter for the present argument. First, it took place in a research setting with no treating clinician, so it is, if anything, a conservative estimate of what clients might provide when a therapist reviews the data. Second, it shows the shape of real longitudinal data: most participants provide most of the data, few provide all of it, and passive streams fill in many days on which self-report is missing. Any system intended for clinical use must therefore handle incomplete series as the normal case.

Measurement embedded in the clinical workflow

When measurement is built into the routine of care, completion can rise sharply. Steidtmann et al. (2025) described the Measurement-Assisted Care (MAC) program, a quality-improvement initiative in an outpatient academic psychiatry service that included 4,665 clients and 176 clinicians over 12 months. Standardized measures, including the Patient Health Questionnaire-9 (PHQ-9) and the Generalized Anxiety Disorder-7 (GAD-7), were added automatically to the electronic check-in process, and responses were routed to clinicians' electronic health record (EHR) inboxes and documentation templates. PHQ-9 completion rose from 5% to 66% of visits in the first rollout phase and from 5% to 60% in the second. Completion after implementation was higher for telehealth visits (70%) than for in-person visits (40%). These are completion rates per visit, not per client, and they describe a single academic service.

Wiebe et al. (2021) reported a related pattern in Canadian outpatient depression and trauma programs. In a pre/post comparison of 18-month periods, the proportion of clients who completed at least one assessment battery at one time point rose from 15.62% (427 of 2,734) with paper administration to 53.98% (1,267 of 2,347) after measures moved to an online portal completed before sessions. The authors noted that this outcome reflects one-time completion and called for research on whether online ROM leads to sustained completion. The finding therefore supports the claim that the mode and timing of collection matter, not the stronger claim that repeated completion is assured.

Human guidance increases completion

Clinical context also supplies something that consumer apps lack: a person. Musiat et al. (2022) meta-analyzed 22 studies, mostly targeting depression and anxiety, that compared guided with unguided web-based mental health interventions. Guidance increased the average amount of intervention completed ($g = 0.29$, 95% CI [0.18, 0.40]) and the proportion of completers (log odds ratio = 0.50, 95% CI [0.34, 0.66]); on average, full completion rates were 12% higher in guided groups. The effects are small, and the studies concerned completion of intervention content rather than measurement. They nevertheless point in the same direction as the review by Lipschitz et al. (2023), which identified integrating human support as one of several intervention-level features that may improve engagement.

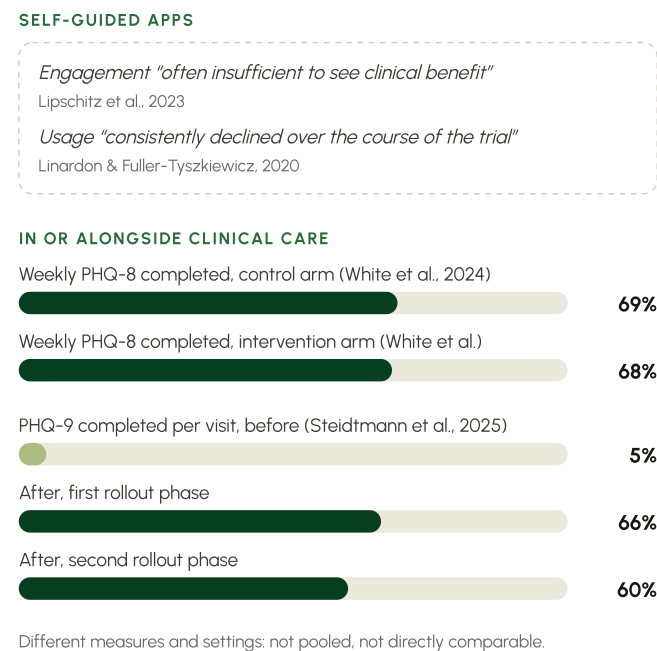


Figure 1. Engagement depends on context. Self-guided apps, in the words of the reviewed authors, against completion reported in or alongside clinical care. The measures differ (share of weekly questionnaires completed; share of visits with a completed questionnaire), so the figures are not pooled and not directly comparable.

Sources: Lipschitz et al., 2023; Linardon & Fuller-Tyszkiewicz, 2020; White et al., 2024; Steidtmann et al., 2025.

What this evidence does and does not show

Taken together, these studies indicate that clients in, or alongside, clinical care provide substantial amounts of repeated self-report and passive data, and that embedding measurement in the workflow and adding human support are associated with more complete data. They do not show that every client will provide data, that completion will be sustained across long treatments, or that any particular level of completion is clinically sufficient. The low retention seen in self-guided settings remains real; it is simply weak evidence about clinical settings, where the conditions that sustain participation are present.

3 Feedback improves outcomes, especially for clients who are not on track

If clients provide data, the next question is whether making those data visible to clinicians benefits clients. The meta-analytic answer since 2020 is a qualified yes.

Rognstad et al. (2023) conducted a three-level meta-analysis of RCTs comparing treatment assisted by measurement feedback systems (MFSs) with treatment as usual for common mental health disorders. Across the included trials, which were conducted mostly in outpatient and university counselling settings, MFSs had a small significant effect on mental health outcomes ($d = 0.14$, 95% CI [0.082, 0.206], $p < .001$). The effect was larger for clients identified as not on track ($d = 0.29$, 95% CI [0.114, 0.464], $p = .003$). The authors noted heterogeneity, few studies for some comparisons, and highly uncertain moderator conclusions.

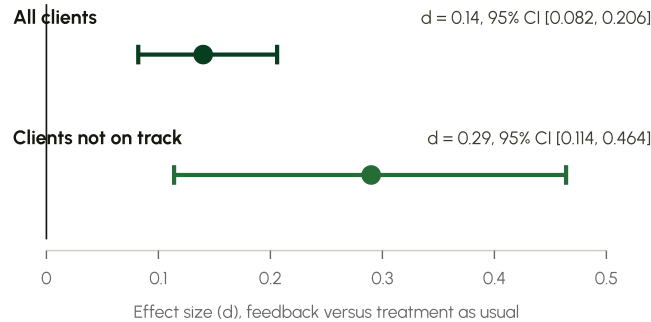


Figure 2. Effects of measurement feedback systems. Effect sizes (d) with 95% confidence intervals from a three-level meta-analysis of randomized trials comparing treatment assisted by measurement feedback systems with treatment as usual. The authors noted heterogeneity and few studies for some comparisons.

Sources: Rognstad et al., 2023.

De Jong et al. (2021) reached a convergent estimate from a broader base of 58 randomized and non-randomized studies (110 effect sizes; 21,699 participants). Progress feedback had a small significant effect on symptom reduction compared with control conditions ($d = 0.15$, 95% CI [0.10, 0.20]), a similar effect for not-on-track cases ($d = 0.17$, 95% CI [0.11, 0.22]), and a small favourable effect on dropout (odds ratio = 1.19, 95% CI [1.03, 1.38]). Barkham et al. (2023), summarizing 11 meta-analytic studies in a narrative review, described overall effects of approximately $g = 0.15$, larger effects when clinical support tools were used for not-on-track cases (approximately $g = 0.36$ to 0.53), and an average success rate difference of approximately 8% over standard care. They also stressed that ROM “does not work for all patients” and that implementation is a major challenge (Barkham et al., 2023, p. 841).

Recent pragmatic studies extend these findings into routine and digital care. Hisler et al. (2025) ran a cluster RCT within NHS Berkshire Talking Therapies, in which therapists supporting internet-delivered cognitive behavioural therapy received feedback from a deep-learning outcome prediction tool. With 51 therapists and 2,394 clients analysed, there was no overall difference between conditions. Among clients flagged as not on track, the feedback condition was associated with a higher likelihood of reliable improvement on the PHQ-9 ($B = 0.17$, $p = .03$), corresponding to roughly a 5% greater likelihood, but not on the GAD-7. Clients in the feedback condition engaged more with the platform, and therapists rated the tool as clear and understandable ($M = 4.30$). The authors were employees of the platform’s provider, and therapists were not blind to condition.

Forand et al. (2025) examined an organization-wide MBC implementation across 755 clinicians and approximately 18,700 clients at a technology-supported psychotherapy practice. In a retrospective, non-randomized pre/post design, client outcomes improved by approximately 5 percentage points from pre- to post-implementation; for the combined PHQ-9 and GAD-7 improvement measure, the post-implementation change was 4.6 percentage points (95% CI [2.8, 6.4]), a relative improvement reported as 23.5% in the abstract. The authors stated that causal attributions are not possible given the design, and they were employees of and held equity in the organization studied.

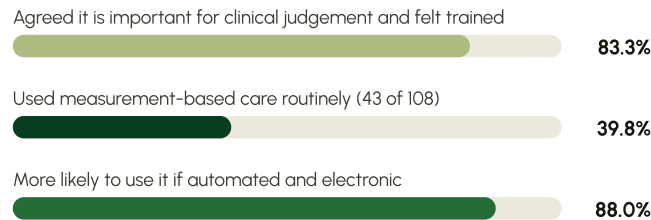
Espel-Huynh et al. (2024) analysed a retrospective cohort of 33,492 adults with clinically elevated anxiety or depression in a blended care program that combined video sessions with asynchronous digital lessons, exercises, and feedback messages sent by providers between sessions. In growth curve models, provider feedback messages in the preceding 7 days were associated with lower GAD-7 scores ($b = -0.12$, $P < .001$) and lower PHQ-9 scores ($b = -0.15$, $P < .001$), with smaller associations for messages 8 to 14 days earlier (GAD-7 $b = -0.07$; PHQ-9 $b = -0.08$). These are unstandardized differences of a fraction of a scale point in an observational, company-authored study, and they should be read as associations rather than effects.

The pattern across these studies is consistent. Feeding longitudinal data back to clinicians is associated with small average benefits, and the benefit appears concentrated among clients whose progress is off track, which is precisely the group for whom a therapist's unaided impression is most likely to be wrong. The evidence does not show that more data produces better outcomes, and it does not show that benefits will appear when data are collected but not used.

4 The usability gap

Clinicians value measurement but use it inconsistently

The weak link between collection and benefit is clinical use. Cheung et al. (2023) surveyed 108 Canadian clinicians and 131 clients with self-reported depression about MBC and technology-enhanced MBC. Most clinicians (83.3%) agreed that MBC is important for clinical judgement and felt trained to interpret scores, yet only 39.8% (43 of 108) used MBC routinely. Most clinicians (88.0%) said they would be more likely to use MBC if it were automated and available electronically, and the authors identified perceptions of efficiency, ease of use, and accessibility as barriers. Most clients (89.3%) felt MBC would be helpful and were willing to use a mobile app to track symptoms. Because recruitment was online, the sample may over-represent people comfortable with digital tools.

108 CANADIAN CLINICIANS

Online recruitment: may over-represent clinicians at ease with digital tools.

Figure 3. The usability gap. Responses of 108 Canadian clinicians in an online survey. Recruitment was online, so the sample may over-represent people comfortable with digital tools.

Sources: Cheung et al., 2023.

Where data are well integrated, clinicians report using them. In the MAC program, more than 90% of surveyed clinicians reported that the data were useful with at least one of their three most recent clients, 51% that they were useful with all three, and 78% that they used the data to guide care (Steidtmann et al., 2025). The contrast between “useful with at least one” and “useful with all three” is itself informative: even in a well-designed program, usefulness varies from client to client, which suggests that the format of the information, not only its availability, determines whether it enters the session.

Collection without integration

Forand et al. (2025) framed the central risk directly: “gathering data without integrating feedback into clinical care does not align with best practices in MBC, and may even be harmful” (Section 1.2). Among their recommendations for implementation was “developing intuitive, user-friendly software platforms that automate key MBC practices and provide decision support” (Forand et al., 2025, Section 5.6). Barkham et al. (2023) similarly reported that clients support ROM “when its purpose is clear and integrated within therapy,” that graphs, more specific feedback, and alerts are helpful, and that organizational, personnel, and resource issues remain the greatest obstacles to adoption (p. 841). The implication is that the same data can be helpful, neutral, or counterproductive depending on how they reach the clinician and whether they are used with the client.

Documentation burden as the context of use

Any new information stream competes with existing demands on clinician time. In a national survey of 1,524 US office-based physicians, respondents reported spending a mean of 1.77 hours per day (95% CI [1.67, 1.87]) completing documentation outside office hours, rising to 1.84 hours among EHR users (Gaffney et al., 2022). That study did not report psychiatry separately, and its data were self-reported. Evidence specific to mental health settings points in the same direction. In a survey at a Canadian mental health organization, 26% (33 of 128) of responding physicians met criteria for burnout on the Mini-Z questionnaire, and 61% (20 of 33) of those attributed their burnout to EHR use; respondents identified chart navigation, information retrieval, and longitudinal medication history among the areas needing improvement (Tajirian et al., 2025). Although these data concern physicians rather than psychotherapists, they describe the environment into which any longitudinal data tool must fit: one in which retrieving and synthesizing information over time is already experienced as a burden.

The usability gap can therefore be stated simply. Clients provide data, and feedback of data can help, but clinicians receive data in forms and at moments that do not fit the work, and they already lack time. A tool that adds one more inbox, one more chart, or one more score to review is unlikely to be used consistently, however good its data.

5 Design principles for clinically useful longitudinal data

The following principles are derived from the evidence above and from the broader literature on interpretable systems in mental health. They are offered as hypotheses for design and evaluation, not as established findings.



Figure 4. From collected to clinically useful. Five of the paper’s seven design principles, in the order a piece of client data meets them. They are offered as hypotheses for design and evaluation, not as established findings.

Sources: The design principles of this paper (1, 2, 3, 5 and 7).

1. The clinician remains in control

Feedback systems support clinical judgement; they do not replace it. The clinician should decide which measures to use, which data streams are relevant for a given client, which observations enter the record, and what is discussed in session. Any summary or flag generated by software should be reviewable, editable, and dismissible, and nothing should be filed to the clinical record without the clinician’s approval. This principle follows from the evidence that benefit depends on clinicians integrating feedback into care (Barkham et al., 2023; Forand et al., 2025): a system that bypasses the clinician removes the mechanism through which feedback appears to help.

2. Interpretable summaries before and within sessions

Raw longitudinal data are not clinically usable on their own. The evidence that graphs, specific feedback, and alerts are helpful (Barkham et al., 2023), together with clinicians’ preference for automated, electronic measurement (Cheung et al., 2023), suggests that the unit of value is a brief, interpretable summary delivered when the clinician can act on it, typically in the minutes before a session and as a reference during it. A useful summary would show change over time against the client’s own baseline, highlight departures from the expected course, and indicate which data are missing, rather than presenting every data point. Because the clearest benefits appear for clients who are not on track (de Jong et al., 2021; Rognstad et al., 2023), surfacing that status early and plainly is likely to matter more than comprehensive dashboards.

3. Traceability of every insight to its source

When software summarizes or interprets data, the clinician must be able to see where each statement came from: which questionnaire item, which day, which journal entry, or which device reading. Joyce et al. (2023) argued that in psychiatry, "explainability" is better approximated by understandability, a function of transparency and interpretability, and that ensuring human understandability of a model's inputs, processes, and outputs is essential for trustworthy systems. Traceability is a practical form of this principle. It allows the clinician to verify a summary against its evidence, to correct it, and to explain it to the client, and it makes errors visible rather than hidden.

4. Association, not causation

Longitudinal client data invite causal stories: poor sleep caused low mood, a difficult meeting caused avoidance. Most patterns detectable in such data are co-occurrences or temporal sequences within one person's series, observed without control of other influences. Joyce et al. (2023) noted that the relationships among symptoms, outcomes, and their determinants in psychiatry are probabilistic. Summaries should therefore use the language of association ("co-occurred with", "preceded", "may relate to") and leave interpretation to the clinician and client. The same discipline applies to evaluating these tools: several of the studies reviewed here are observational, and their authors were careful to describe associations (Espel-Huynh et al., 2024; Forand et al., 2025).

5. Minimal client burden

Incomplete data are the normal case (White et al., 2024). Systems should ask for the smallest amount of active input that serves a clinical purpose, make the purpose visible to the client, and use passive or already-available data where clients choose to share them. Barkham et al. (2023) noted that greater frequency of data collection matters more for shorter-term therapies, which implies that frequency should be set by clinical need rather than by what the technology allows. Interpretations should tolerate gaps and show them, rather than imputing values silently.

6. Privacy by design and client choice

Longitudinal data about mood, sleep, relationships, and daily activity are among the most sensitive data a person can share. Clients should choose what they share, with whom, and at what level of detail, for example per device and per type of data, and should be able to change those choices. Data should be minimized to what serves care, protected in storage and transit, and governed by the privacy laws of the client's jurisdiction. Privacy is also a precondition for data quality: clients

who do not trust how their information will be used have reason to withhold or edit it. We note that this principle rests on ethical and regulatory reasoning rather than on outcome evidence reviewed here.

7. Measure usefulness, not only engagement

Most evaluations of digital tools report engagement or completion. For tools intended to support clinicians, the outcome that matters is whether the information changed what happened in care. Steidtmann et al. (2025) offer a practical model by asking clinicians whether the data were useful with specific recent clients and whether they used the data to guide care. Lipschitz et al. (2023) proposed a minimum engagement reporting framework for digital interventions; an analogous framework for clinical usefulness might include the proportion of sessions in which data were reviewed, the proportion in which they were discussed with the client, the time clinicians spend reviewing them, the rate at which summaries are corrected or rejected, and client outcomes, particularly for those who are not on track. Measuring correction and rejection rates is especially important for systems that generate interpretations, because it directly indexes whether clinicians find them trustworthy.

An example of ongoing work

Several organizations are developing systems along these lines. One of them is Cognitai Health Inc., the company of which the Cognitai Research Department is the research function, which is building clinician-facing software for organizing information shared between sessions; we mention it only as an example of ongoing work, and no performance data are reported or implied.

6 Limitations

The evidence reviewed here has important limitations. First, it is heterogeneous. The studies differ in population (adults with major depressive disorder, mixed outpatient samples, clients of internet-delivered and blended care), in setting (research platforms, academic psychiatry, national health services, employer-sponsored and commercial services), and in what was measured (per-visit completion, one-time completion, proportion of weekly questionnaires, intervention completion). We have reported each statistic as defined by its authors, but they cannot be pooled or directly compared.

Second, the effects of feedback are small on average. Overall effect sizes of approximately $d = 0.14$ to 0.15 (de Jong et al., 2021; Rognstad et al., 2023) are clinically meaningful at scale but modest for any individual client, and estimates for not-on-track clients vary across meta-analyses and depend on whether clinical support tools were used (Barkham et al., 2023).

Third, several key studies are observational or non-randomized. The completion gains reported by Steidtmann et al. (2025) and Wiebe et al. (2021) are pre/post comparisons within single organizations, and the outcome associations reported by Espel-Huynh et al. (2024) and Forand et al. (2025) come from retrospective designs that preclude causal inference.

Fourth, several studies were authored by employees of the companies whose services or platforms were evaluated (Espel-Huynh et al., 2024; Forand et al., 2025; Hisler et al., 2025). This does not invalidate their findings, but it warrants caution and independent replication. We, too, have a conflict of interest, disclosed in the Disclosure section.

Fifth, the evidence on documentation burden comes from physician samples, not psychotherapists (Gaffney et al., 2022; Tajirian et al., 2025), and the survey evidence on clinician attitudes may over-represent clinicians already favourable to technology (Cheung et al., 2023).

Sixth, this is a narrative review. We did not conduct a systematic search, dual screening, or formal risk-of-bias assessment, and the restriction to sources published from 2020 onward excludes earlier foundational work on feedback-informed treatment and homework in psychotherapy. The design principles we propose are derived from this evidence but have not themselves been tested.

7 Conclusion

The claim that clients will not provide data between sessions is drawn largely from settings without a clinician. In clinical care, where a therapist expects the data and a session gives it purpose, clients provide substantial repeated self-report and passive data, especially when measurement is embedded in the workflow and supported by people. Feeding those data back to clinicians is associated with small average benefits and larger benefits for clients who are not on track. The constraint that now limits these benefits is less the willingness of clients to share than the ability of clinicians to use what is shared: to see it at the right moment, in a form they can interpret quickly, trace to its source, and discuss with the client, without adding to a documentation burden that is already heavy. Research and design effort should shift accordingly, from maximizing collection toward making longitudinal data clinically useful, and evaluations should measure that usefulness directly.

References

- Barkham, M., de Jong, K., Delgadillo, J., & Lutz, W. (2023). Routine outcome monitoring (ROM) and feedback: Research review and recommendations. *Psychotherapy Research*, 33(7), 841–855. <https://doi.org/10.1080/10503307.2023.2181114>
- Cheung, B. S. W., Murphy, J. K., Michalak, E. E., Liu, J., Yang, X., Wang, X., Chen, J., & Lam, R. W. (2023). Barriers and facilitators to technology-enhanced measurement based care for depression among Canadian clinicians and patients: Results of an online survey. *Journal of Affective Disorders*, 320, 1–6. <https://doi.org/10.1016/j.jad.2022.09.055>
- de Jong, K., Conijn, J. M., Gallagher, R. A. V., Reshetnikova, A. S., Heij, M., & Lutz, M. C. (2021). Using progress feedback to improve outcomes and reduce drop-out, treatment duration, and deterioration: A multilevel meta-analysis. *Clinical Psychology Review*, 85, Article 102002. <https://doi.org/10.1016/j.cpr.2021.102002>
- Espel-Huynh, H. M., Wang, L., Lattie, E. G., Wickham, R. E., Varra, A., Chen, C. E., Lungu, A., & Lee, J. L. (2024). Clinical effects of asynchronous provider-guided practice sessions during blended care therapy for anxiety and depression: Pragmatic retrospective cohort study. *Journal of Medical Internet Research*, 26, Article e60502. <https://doi.org/10.2196/60502>
- Forand, N. R., Nettiksimmons, J., Brownell, A., Anton, M. T., Truxson, R., Green, B., & Marshall, C. (2025). The impact of measurement based care at scale: Examining the effects of implementation on patient outcomes and provider behaviors. *Frontiers in Health Services*, 5, Article 1659238. <https://doi.org/10.3389/frhs.2025.1659238>
- Gaffney, A., Woolhandler, S., Cai, C., Bor, D., Himmelstein, J., McCormick, D., & Himmelstein, D. U. (2022). Medical documentation burden among US office-based physicians in 2019: A national study. *JAMA Internal Medicine*, 182(5), 564–566. <https://doi.org/10.1001/jamainternmed.2022.0372>
- Hisler, G. C., Young, K. S., Cumpasasoiu, D. C., Palacios, J. E., Duffy, D., Enrique, A., Keegan, D., & Richards, D. (2025). Incorporating a deep-learning client outcome prediction tool as feedback in supported internet-delivered cognitive behavioural therapy for depression and anxiety: A randomised controlled trial within routine clinical practice. *Counselling and Psychotherapy Research*, 25(1), Article e12771. <https://doi.org/10.1002/capr.12771>

- Joyce, D. W., Kormilitzin, A., Smith, K. A., & Cipriani, A. (2023). Explainable artificial intelligence for mental health through transparency and interpretability for understandability. *npj Digital Medicine*, 6, Article 6. <https://doi.org/10.1038/s41746-023-00751-9>
- Linardon, J., & Fuller-Tyszkiewicz, M. (2020). Attrition and adherence in smartphone-delivered interventions for mental health problems: A systematic and meta-analytic review. *Journal of Consulting and Clinical Psychology*, 88(1), 1–13. <https://doi.org/10.1037/ccp0000459>
- Lipschitz, J. M., Pike, C. K., Hogan, T. P., Murphy, S. A., & Burdick, K. E. (2023). The engagement problem: A review of engagement with digital mental health interventions and recommendations for a path forward. *Current Treatment Options in Psychiatry*, 10(3), 119–135. <https://doi.org/10.1007/s40501-023-00297-3>
- Musiat, P., Johnson, C., Atkinson, M., Wilksch, S., & Wade, T. (2022). Impact of guidance on intervention adherence in computerised interventions for mental health problems: A meta-analysis. *Psychological Medicine*, 52(2), 229–240. <https://doi.org/10.1017/S0033291721004621>
- Rognstad, K., Wentzel-Larsen, T., Neumer, S.-P., & Kjøbli, J. (2023). A systematic review and meta-analysis of measurement feedback systems in treatment for common mental health disorders. *Administration and Policy in Mental Health and Mental Health Services Research*, 50(2), 269–282. <https://doi.org/10.1007/s10488-022-01236-9>
- Steidtmann, D., Raffel, K. E., Green, J., MacPhee, E., Nagle-Yang, S., Schwenk, S., Cooke, D., Santisteban, A. C., & Dempsey, A. G. (2025). Using technology to support measurement-based care: Design and implementation of the Measurement-Assisted Care program. *Psychiatric Services*, 76(2), 126–132. <https://doi.org/10.1176/appi.ps.20240135>
- Tajirian, T., Lo, B., Strudwick, G., Tasca, A., Kendell, E., Poynter, B., Kumar, S., Chang, P. B., Kung, C., Schachter, D., Zai, G., Kiang, M., Hoppe, T., Ling, S., Haider, U., Rabel, K., Coombe, N., Jankowicz, D., & Sockalingam, S. (2025). Assessing the impact on electronic health record burden after five years of physician engagement in a Canadian mental health organization: Mixed-methods study. *JMIR Human Factors*, 12, Article e65656. <https://doi.org/10.2196/65656>
- White, K. M., Carr, E., Leightley, D., Matcham, F., Conde, P., Ranjan, Y., Simblett, S., Dawe-Lane, E., Williams, L., Henderson, C., & Hotopf, M. (2024). Engagement with a remote symptom-tracking platform among participants with major depressive disorder: Randomized controlled trial. *JMIR mHealth and uHealth*, 12, Article e44214. <https://doi.org/10.2196/44214>

Wiebe, D. E., Remers, S., Nippak, P., & Meyer, J. (2021). Evaluation of an online system for routine outcome monitoring: Cross-sectional survey study. *JMIR Mental Health*, 8(12), Article e29243. <https://doi.org/10.2196/29243>