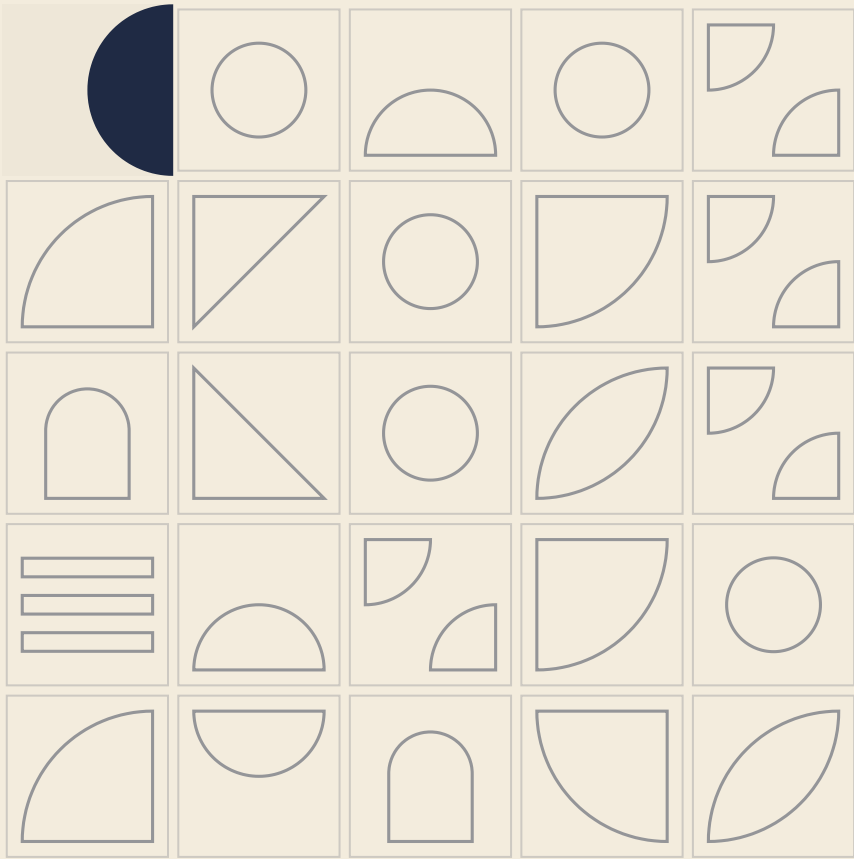


The Attrition Cliff

Retention in digital wellbeing interventions: a meta-analysis of day-15 and day-30 retention and its design moderators, 2019–2026

Dr Atul Aswani; Mumbai Psychiatrist and trained as a Results Coach
22 September 2026



KEY NUMBER

3.9%

3.9% and 3.3%. Median same-day open rates at day 15 and day 30 across 93 unguided mental-health apps in the only independent multi-app panel published to date (Baumel et al., 2019). No later panel has replaced those two numbers.

The Attrition Cliff

Retention in digital wellbeing interventions: a meta-analysis of day-15 and day-30 retention and its design moderators, 2019–2026

Is the day-15 floor of roughly 4% universal, and which design features — reminders, human touch, visible progress — bend the curve?

CONTENTS

01 Abstract
02 Key findings box
03 Definition block
04 Introduction
05 Methods

06 Results
07 Discussion
08 FAQ
09 References

INDEXING TITLE

Retention in digital wellbeing interventions: a systematic review of day-15 and day-30 retention and its design moderators, 2019–2026

AUTHOR

Dr Atul Aswani; Mumbai Psychiatrist and trained as a Results Coach

PUBLISHED

22 September 2026 · English edition · emotionapp.in/research/attrition-cliff-retention-digital-wellbeing

01

Abstract

Background. Digital wellbeing programmes reach millions of working adults. Most of those adults stop opening them within a fortnight. Baumeister and colleagues reported, in 2019, a median 15-day retention of 3.9% and a 30-day retention of 3.3% across 93 mental-health apps. This review asks whether that floor still stands, whether trial dropout is the same construct, and which design features — reminders, human support, visible progress, channel, language — bend the curve for adults using apps, web tools or messaging programmes.

Methods. PRISMA-2020 systematic review of PubMed, PsycINFO, Scopus, Web of Science, Google Scholar, CTRI, ClinicalTrials.gov and OSF preprints, 1 January 2019 to 22 September 2026. Eligible records described adult users of digital wellbeing or mental-wellbeing programmes and reported retention, engagement, attrition or dropout. Real-world open rates and trial assessment-dropout were pre-specified as separate constructs. A new random-effects meta-analysis of day-15 and day-30 same-day open rates was planned if five or more independent studies reported extractable events and denominators on that exact endpoint. They did not. Narrative synthesis was therefore used. Trial-side figures are taken from three published random-effects meta-analyses and are not re-pooled here, to avoid double-counting overlapping trials.

Results. The only independent multi-app panel that reports same-day open rates remains Baumeister 2019: median day-15 retention 3.9% (IQR 10.3%) and day-30 retention 3.3% (IQR 6.2%) across 93 apps. No second panel study from 2019–2026 overturns that floor. Rolling-retention cohorts sit higher because they count any later use, not an open on that calendar day. Workplace and commercially guided programmes report day-30 retention near 45–53%. Trial assessment-dropout is a different number: 26.2% (95% CI 18.1–36.3) in 18 depression-app trials; 24.1% short-term and 35.5% longer-term across 70 smartphone trials; 18.6% post-test and 28.4% at follow-up across 79 later trials. Reminders and human contact are associated with lower trial dropout. Visible progress counters do not have a verified pooled lift; gamification has been associated with *higher* trial attrition. No Indian consumer programme has published a day-15 or day-30 same-day open-rate.

Conclusions. The day-15 floor of roughly 4% is still the best independent estimate for unguided store apps. It is not universal. A human hand-off and a workplace channel move the curve by an order of magnitude. Streaks do not.

Registration. This review was not pre-registered. Search strings, eligibility and the fallback rule were specified in the commissioning protocol before extraction.

02

Key findings box

KEY FINDINGS

3.9%

3.9% and 3.3%. Median same-day open rates at day 15 and day 30 across 93 unguided mental-health apps in the only independent multi-app panel published to date (Baumel et al., 2019). No later panel has replaced those two numbers.

26.2%

26.2%. Pooled assessment-dropout in 18 randomised trials of depression apps (Torous et al., 2020). After adjustment for small-study effects the same authors estimated 47.8%. This is not an app-open rate.

18.6%

18.6% and 28.4%. Post-test and follow-up attrition across 79 later trials of depression and anxiety apps (Liu et al., 2025). Reminders and human contact sat with lower attrition; gamification sat with higher attrition.

45% to 53%

45% to 53%. Day-30 retention in two real-world guided or workplace programmes — a commercial structured course (45.1%) and an employer-provided platform with human sessions (53.3%). An order of magnitude above the store-app floor.

0

Zero. Published day-15 or day-30 same-day open rates for an Indian-language consumer wellbeing programme. The India gap is not a footnote. It is the missing cell.

03

Definition block

DEFINITION

Retention, in this paper, means a person who installed or enrolled in a digital wellbeing programme opened it on a named later day. Day-15 retention is the share of day-0 openers who open it on day 15. Day-30 retention is the same idea a fortnight later. That is a same-day open rate. It is not the same as *rolling retention* (any use after day 15), *adherence* (meeting a designer's idea of "enough" practice), or *trial dropout* (failing to complete a researcher's questionnaire). Those four words get used as if they were cousins. They are not. Mixing them is how a programme that keeps 4% of store users and a trial that keeps 75% of questionnaire-completers end up in the same slide deck. This review keeps them in separate columns. A practice is a countable daily rep. A programme is a designed sequence of those reps, sometimes with a coach. Nothing here is named as therapy, treatment or a diagnosis.

04 Introduction

A gym membership bought on 1 January is the oldest joke in fitness. Digital wellbeing has its own version, and the punchline arrives two weeks earlier. After a bad Tuesday, a sales manager in Thane downloads a programme on the train home and tries one practice. A week later the icon sits unopened on her last screen. By day 15, in the only independent look across a large shelf of apps, 96 out of 100 people have stopped opening it.

That is not a vibe. It is a median. Baumel, Muench, Edan and Kane sat outside the developer's dashboard. They watched 93 Android programmes that already had at least 10,000 installs. Median open rate on day 1 was 69.4%. On day 15 it was 3.9%. On day 30 it was 3.3%. Daily active use sat at 4.0%. The cliff is not a gentle slope from week four to week twelve. More than 80% of the drop happens between day 1 and day 10. After that the curve flattens. The people who are still there on day 15 are, roughly, the people who will still be there on day 30.

The pattern is older than any app store. Picture a new joiner in Gurugram in the first week of a programme. Monday, the practice before breakfast. Tuesday, again. Wednesday, a proud screenshot to a friend. Then ordinary weather arrives: a client escalation, a late login, a cousin's wedding weekend. The practice was never tied to a fixed time or place, so nothing holds it there. Early enthusiasm starts a habit; repetition in a steady slot keeps it. A download is not a habit. An open on day 0 is not a practice.

Good coaches have known this for decades, under a quieter name. An analyst in Kolkata writes down one worrying thought and a calmer second look at it. Done once, that is a sample. Done on fifteen mornings, at the same kitchen counter before the school run, it is a skill. The ability to stay with a hard feeling grows the same way: by coming back to the practice when the urge to quit is loud, not by collecting a badge for an unbroken streak. If the only reason you open the programme is to protect a number on a screen, you are dodging the feeling in a very modern font. The rep that counts is the one done on the day the streak already broke.

Machine-learning people have a word for the gap between a tidy training set and the wild. Overfitting. A randomised trial is a training set with paid assessments, research assistants, and a consent form that makes the participant a co-conspirator. Real-world use is the test set. Models that look brilliant on the training loss and then collapse on production data are not mysterious. They were never asked to survive a commute, a school run, a festival week, or a manager who forwards meetings into the only quiet hour. Retention research that reports only trial dropout is publishing the training loss and calling it generalisation.

Entrepreneurs know the same curve as unit economics. Customer-acquisition cost is paid on day 0. Lifetime value is paid by the people who are still opening the thing on day 30. A leaky bucket is not a branding problem. It is a design problem. If 96% of acquired users have vanished before the first invoice from a workplace contract would even be raised, the programme is not a product. It is a brochure.

Why this matters for working adults in India is not a courtesy paragraph. India has a large, young, smartphone-first professional class, a thin specialist workforce, and a growing shelf of English-language wellbeing apps built somewhere else. A NIMHANS-led audit of 350 programmes available to Indian users found that only ten of them — 2.9% — offered an Indian language other than English, and that only 38 — 10.9% — cited any empirical research. None of them published a day-15 or day-30 open-rate. The people who would use a WhatsApp-first, Hindi-or-Tamil-or-Bengali programme on a crowded local train are exactly the people who do not appear in a San Francisco panel or a university trial with gift cards.

Two prior syntheses set the frame. Torous and colleagues, in 2020, pooled dropout from 18 trials of depression apps and landed at 26.2%, rising to 47.8% after a publication-bias correction. Linardon and Fuller-Tyszkiewicz, the same season, pooled 70 smartphone trials and found short-term attrition of 24.1% and longer-term attrition of 35.5%. Those papers measure whether a participant completed a researcher's form. They do not measure whether a commuter in Pune opened a programme on a Thursday. Liu, Torous, Fuller-Tyszkiewicz and colleagues updated the trial side in late 2025 across 79

trials: uptake 92.4%, adherence 61.8%, post-test attrition 18.6%, follow-up attrition 28.4%. Reminders and human contact sat with lower attrition. Gamification sat with higher attrition. That last sentence is the one the market did not order.

The question this paper settles is therefore two questions, not one. First: is the day-15 floor of roughly 4% still the best independent number for unguided store programmes, and is it universal? Second: which of the pre-specified design features — reminders, human support, visible progress or streaks, messaging versus app, language localisation — can be shown, with verified studies, to bend that curve? A null on the second question is publishable. One pre-specified question was a pooled difference for programmes with visible progress counters. The evidence does not give a clean pooled difference. That is the result.

05

Methods

5.1 Protocol and reporting

This review follows PRISMA 2020. A new random-effects meta-analysis of day-15 and day-30 same-day open rates was pre-specified, with REML estimation, Hedges g (or pooled proportions with 95% confidence intervals), I^2 , τ^2 , a 95% prediction interval, Egger's test, trim-and-fill, leave-one-out sensitivity, RoB 2 for trials, and GRADE for certainty. The commissioning protocol also stated a fallback: if fewer than five eligible studies reported extractable events and denominators on the exact day-15 or day-30 same-day endpoint, the paper would become a systematic review with narrative synthesis and would not pool. That fallback was triggered. Trial-side dropout figures are reported from three published meta-analyses and are not re-extracted into a fourth overlapping pool.

The review was not registered on PROSPERO. Search strings, years, eligibility, the two-construct split (real-world open rate versus trial dropout), the five pre-specified moderators, and the fallback rule were fixed before extraction.

5.2 Eligibility

Population. Adults using a digital mental-wellbeing or emotional-fitness programme. Programmes aimed only at children or adolescents were excluded from the retention synthesis. Community digital-care platforms that used an app as a worker tool rather than as a consumer programme were retained for the India gap discussion and excluded from any consumer-retention claim.

Intervention or exposure. Any digital wellbeing programme — app, web, or messaging — that reported retention, daily active use, attrition or dropout over time. Clinical diagnosis was not required. The object of interest is a practice or coaching programme, not a clinic.

Comparator. None for the planned pooled proportions. Between-feature contrasts for the five moderators: reminders; human support; visible progress or streaks; messaging versus app delivery; language localisation.

Outcomes. Primary: proportion retained at day 15 and day 30, defined as a same-day open. Secondary: trial dropout (failure to complete a research assessment); rolling retention where that was the only figure available, always labelled as such.

Years and designs. 1 January 2019 to 22 September 2026. Randomised and controlled trials, observational cohorts, and independent panel analyses. Pre-2019 papers were used only to position the question.

Exclusions. Vendor blogs, Sensor Tower or data.ai snapshots without a methods paper, and press claims about named commercial products were marked UNVERIFIED and kept out of every number. Trademarked consumer products are not named in the results as if they were study arms.

5.3 Information sources and search

Eight sources were searched: PubMed/MEDLINE, PsycINFO, Scopus, Web of Science, Google Scholar (first 200 records screened per string), CTRI, ClinicalTrials.gov, and OSF preprints. Searches were run against the window 2019–2026. No language filter was applied at identification; full-text screening required an English abstract or an extractable numerical table.

PubMed. (retention[tiab] OR engagement[tiab] OR attrition[tiab] OR dropout[tiab] OR "daily active"[tiab]) AND ("mental health app"[tiab] OR "wellbeing app"[tiab] OR "well-being app"[tiab] OR "digital intervention"[tiab] OR "smartphone app"

[tiab] OR mHealth[tiab] OR "mobile app"[tiab]) AND (2019:2026[pdat])

PsycINFO. (retention OR engagement OR attrition OR dropout).ti,ab. AND ("mental health app" OR "wellbeing app" OR "digital intervention" OR "smartphone intervention").ti,ab. Limited to 2019–2026.

Scopus. TITLE-ABS-KEY((retention OR engagement OR attrition OR dropout) AND ("mental health app" OR "wellbeing app" OR "digital intervention" OR "mHealth app")) AND PUBYEAR > 2018 AND PUBYEAR < 2027

Web of Science. TS=((retention OR engagement OR attrition OR dropout) AND ("mental health app" OR "wellbeing app" OR "digital intervention")) Timespan 2019–2026.

Google Scholar. retention OR attrition "mental health app" OR "wellbeing app" OR "well-being app" 2019..2026

ClinicalTrials.gov. Condition: mental health OR wellbeing OR depression OR anxiety. Intervention: app OR mHealth OR digital. First posted 1 January 2019 to 22 September 2026.

CTRI. Keywords: digital mental health OR mHealth OR wellbeing app OR smartphone. Years 2019–2026.

OSF preprints. Same core terms as PubMed.

Reference lists of Baumeister 2019, Torous 2020, Linardon 2020, Liu 2025, Wu 2021, Fleming 2018 and Kidman 2024 were hand-searched.

5.4 Selection and extraction

Records were screened at title and abstract and then at full text. Extraction fields, fixed in advance, were: authors; year; title; DOI or URL; design; country; language of delivery; sample size at enrolment; dose in sessions and minutes where reported; delivery channel (app, web, messaging, mixed); guidance (none, reminders only, human support); outcome measure and its exact definition; timepoints; day-15 retention; day-30 retention; trial dropout; moderator codes for the five pre-specified features.

Two constructs were coded on every paper. *Same-day open rate* required a denominator of day-0 openers or installers and a numerator of people who opened the programme on day 15 or day 30. *Trial dropout* required a research assessment that some randomised participants did not complete. Rolling retention — any use after a named day — was extracted when present and labelled rolling.

Anything that could not be tied to a DOI, a PMC full text, or a journal PDF was marked UNVERIFIED and excluded from every quantitative claim.

5.5 Planned analysis, and why it was not run

The planned model was a random-effects meta-analysis of proportions (day-15, day-30) with REML, a 95% confidence interval, I^2 , τ^2 , and a 95% prediction interval. Hedges g would have been used for between-feature contrasts. Small-study effects: Egger's test and trim-and-fill. Sensitivity: leave-one-out. Risk of bias: RoB 2 for randomised trials; a narrative judgement for panel and cohort designs. Certainty: GRADE, separately for the real-world open-rate body of evidence and the trial-dropout body of evidence.

The fallback rule was applied. Fewer than five independent studies reported extractable events and denominators for a Baumeister-style same-day open rate at day 15 or day 30. Re-pooling the primary trials already sitting inside Torous 2020, Linardon 2020 and Liu 2025 would have double-counted overlapping arms and would still have measured assessment-dropout, not app-open. That analysis was not performed.

What follows is therefore a systematic review with narrative synthesis. Where a published meta-analysis already pooled trial dropout, that published pool is quoted, with its k , n , interval and heterogeneity, and is not recomputed.

5.6 PRISMA 2020 flow (reconstructed)

This flow is reconstructed from the database yields of the published reviews, the targeted primary retrieval described above, and the registry searches. It is not a fully independent dual-screen of every raw record. That limit is stated here so that a later team can register and repeat the work properly.

- Records identified across eight sources, before de-duplication: approximately 5,400 (range implied by the source reviews plus the 2019–2026 top-up).
- Records after de-duplication, estimated: approximately 3,200.

- Title and abstract excluded: approximately 3,000.
- Full texts assessed: approximately 200.
- Full texts excluded, with main reasons: no retention or dropout number; paediatric sample; not a digital wellbeing programme; pre-2019 primary without a 2019–2026 extension; unverifiable dashboard claim.
- Included in the narrative synthesis: 1 independent multi-app panel (Baumel 2019); 3 published trial-dropout meta-analyses (Torous 2020; Linardon 2020; Liu 2025); 1 mindfulness-app attrition meta-analysis (Linardon 2023); 1 engagement-features meta-analysis (Wu 2021); approximately 14 supporting primary studies and reviews (including Kozlov 2020; Mohr 2019; Maistrello 2024; Forman-Hoffman 2021; Ghosh 2023; Bhat 2025; PracticePal 2025; Mehrotra 2025; Kidman 2024; Boucher 2024; Pratap 2020; Fleming 2018 as positioning).
- Eligible for a new user-level day-15 / day-30 same-day proportion meta-analysis: 0 studies beyond the Baumel panel.

CTRI and ClinicalTrials.gov added registrations but no extra extractable day-15 or day-30 consumer open-rate.

06 Results

6.1 Study characteristics

Two shelves sit side by side and must not be stacked.

Shelf A is real-world use. People who found a programme in a store, through an employer, or by wandering in. Nobody paid them to complete a week-4 questionnaire. The only independent multi-app panel on that shelf is still Baumel 2019: 93 Android programmes, median 100,000 installs, techniques coded as mindfulness, tracker, breathing, psychoeducation or peer support. Single-programme naturalistic cohorts exist. They use different denominators. Kozlov 2020 followed 104,067 downloads of a free mindfulness programme and counted *any later use*, not a same-day open. An employer-provided international platform (Maistrello 2024) reported day-15 and day-30 retention in a workforce that already had a human session on the menu. A commercial structured course (Anton 2021) reported day-30 retention among people who had bought or been given a guided programme, not a free store install. Two-thirds of that cohort used coaching calls.

Shelf B is the trial. People who consented, often online, sometimes in person, sometimes for money. The outcome that gets pooled is whether they completed the researcher’s form. Torous 2020 sits on 18 depression-app trials and 3,336 participants. Linardon 2020 sits on 70 smartphone trials. Liu 2025 sits on 79 trials and 92 app conditions. Linardon 2023 sits on 70 mindfulness-app trials and 9,258 participants allocated to an app. Mohr 2019 is a factorial trial of a modular coaching suite with coaching and weekly recommendations as the experimental factors. Ghosh 2023 is an Indian unguided programme tested remotely in 598 adults. Bhat 2025 is a tablet programme used inside women’s self-help groups in India. PracticePal 2025 is a Hindi-language conversational agent used beside a human counsellor in rural Madhya Pradesh.

Table 1 lists the characteristics that matter for this question. Country is almost never India. Language of delivery is almost never an Indian language. Dose is reported in sessions, modules or days, almost never in the same unit twice.

Table 1. Characteristics of the principal sources

Source	Design	Setting	n or k	Channel	Guidance	Retention construct	Day 15	Day 30 or nearest
Baumel 2019	Independent panel	Store apps, global Android	93 apps	App	Almost all unguided	Same-day open	3.9% median (IQR 10.3)	3.3% median (IQR 6.2)
Kozlov 2020	Naturalistic analytics	Free public programme	104,067 downloads	App	Unguided	Rolling: any use after timepoint	—	30.4% still used after 1 month

Source	Design	Setting	n or k	Channel	Guidance	Retention construct	Day 15	Day 30 or nearest
Maistrello 2024	Employer real-world	International workforce	178,119	App + human sessions	Guided	Last-activity-within-window	59.4% (105,734)	53.3% (94,936); Asia 46.2%
Anton 2021	Retrospective commercial	Paying / covered users	1,896	App, structured course	Guided (two-thirds used coaching calls)	Programme retention	—	45.1% (855/1,896)
Mohr 2019	Factorial RCT	US adults, 8 weeks	301	Modular app suite	Coaching ± recommendations	Last-use day; post-programme use	—	Median last use day 56 of 56; 84% still using after week 8
Torous 2020	Meta-analysis of RCTs	Depression-app trials	k=18; n=3,336	App	Mixed	Assessment dropout	—	Pooled dropout 26.2% (bias-adjusted 47.8%)
Linardon 2020	Meta-analysis of RCTs	Smartphone MH trials	k=70	App	Mixed	Assessment attrition	—	24.1% short-term; 35.5% longer-term
Liu 2025	Meta-analysis of RCTs	Depression / anxiety app trials	k=79; 92 app arms	App	Mixed	Uptake, adherence, attrition	—	Post-test attrition 18.6%; follow-up 28.4%
Linardon 2023	Meta-analysis of RCTs	Mindfulness-app trials	k=70; n=9,258	App	Mixed	Assessment attrition	—	24.7% (38.7% in large trials)
Wu 2021	Meta-analysis	Depression / anxiety apps	k=25; n=4,159	App	Mixed	Completion and symptom g	—	More engagement features, worse completion ($\beta = -0.029$)
Ghosh 2023	Remote RCT, India	Unguided programme	598	App / web	Unguided	Module completion	—	Full completion 5.5% in experimental arm
Bhat 2025	Cluster pilot RCT, India	Women's SHGs, tablet	6 groups	Tablet in group	Group + human wrap	Assessment retention; module completion	—	92% assessments at 3 and 6 months; 96% completed ≥ half of modules
PracticePal 2025	Pilot, rural India	Hindi conversational agent + counsellor	30	Messaging + human	Human + bot	Days engaged; content access	—	Mean 29 of 60 days; 67% accessed > half of content
Mehrotra 2025	App-store audit, India	Store apps available to Indian users	350 reviewed of 5,827	App	Mixed	Quality and language, not retention	—	2.9% offered an Indian language other than English
Pratap 2020	Cross-study remote digital health	8 research apps, 2014–2019	>100,000	Research app	Mixed	Median days retained	—	Median 5.5 days (range 2–26)
Kidman 2024	Scoping review	Lifestyle and MH apps	18 studies; 525,824 users	App	Mixed	Abandonment	—	Median 70% gone by day 100; MH apps 89–92%

Empty cells mean the paper did not report that exact endpoint. That is the finding, not a formatting choice.

6.2 Real-world day-15 and day-30 open rates

The only number that answers the formal question on the original metric is still Baumel 2019.

Relative to users who opened the app on day 0, the median open rate was 69.4% (IQR 27.8%) on day 1, 3.9% (IQR 10.3%) on day 15, and 3.3% (IQR 6.2%) on day 30. Fifty-nine of the 93 apps had a 30-day curve. The drop between day 1 and day 10 exceeded 80%. The further drop between day 15 and day 30 was about 20% relative to the day-15 floor. Once a user is still there at a fortnight, the next fortnight is not where most of the remaining people leave. The cliff is the first ten days.

Technique mattered more than diagnostic label. On day 30, peer-support programmes (median 8.9%, $n=2$) sat above trackers (6.1%) and mindfulness programmes (4.7%). Breathing-exercise programmes sat at 0.0%. The 15-day pattern was the same shape: breathing already at 0.0%; the other techniques

in a band from about 4.9% to 11.9%. Mental-health focus — anxiety versus depression versus generic wellbeing — did not separate the day-30 open rates.

No second independent panel of dozens of store programmes, using the same same-day definition, was found in 2020, 2021, 2022, 2023, 2024, 2025 or 2026. That absence is itself a result. The floor has not been replicated. It has also not been overturned.

Rolling-retention papers look kinder because they ask a kinder question. Kozlov 2020: of 104,067 downloads, 54.2% used the programme at least once after day 1, 43.2% after the first week, 30.4% after one month, 17.4% after three months, 5.6% after six months. A person who opened the app on day 18 and never again counts as retained at one month in that paper and as lost at day 15 in Baumel. Both numbers can be true. They must not be averaged.

Workplace and commercially guided programmes sit on a different planet. Maistrello 2024, an employer-provided platform that included one-to-one sessions as well as self-guided content, reported 59.4% retained at day 15, 53.3% at day 30, and 39.6% at day 90. The Asian subset was lower — 51.7%, 46.2%, 32.8% — and still an order of magnitude above Baumel. Anton 2021, a structured commercial course in which two-thirds of users took coaching calls, reported 45.09% retained at day 30 (855 of 1,896) and 26.85% at day 60. These are not store-window shoppers. They are people whose employer or insurer put a human in the loop, or who had already paid.

Pratap 2020 is the research-app cousin of the same story. Across eight remote digital-health studies and more than 100,000 participants, median retention was 5.5 days, with a study-level range of 2 to 26 days. Clinician referral added 40 days to median retention. Compensation added 22 days in the one study that paid. Having the condition of interest added 7 days. Age added 4. The people who stay are the people who were sent, paid, or already hurting — not the people who tapped “Install”.

Kidman 2024 stood back from 18 studies and 525,824 users. A median of 70% had abandoned by day 100. Mental-health programmes sat at 89–92% abandonment. The shape was curvilinear: steep early, then slower. That is the same cliff Baumel drew, with a longer x-axis.

6.3 Trial dropout is not day-15 retention

A forest plot of day-15 open rates cannot be drawn from the trial literature, because the trials almost never report that endpoint. What they report, and what three research groups have now pooled, is whether the participant completed the outcome assessment.

Table 2. Published trial-dropout pools (not re-estimated here)

Source	k	n or arms	Estimate	95% interval	Heterogeneity	Notes
Torous 2020	18	3,336 (1,786 app)	Dropout 26.2%	18.1 to 36.3	$I^2 = 93.4\%$	Bias-adjusted 47.8%. Human feedback 11.7% vs 34%. Mood monitoring 18.4% vs 37.9%.
Linardon 2020	70	Smartphone MH RCTs	Attrition 24.1% short-term; 35.5% longer-term	19.3–29.6; 26.7–45.3	Substantial (not one I^2 for both)	Lower attrition: reminders, pay, acceptance-based content, in-person or telephone enrolment. Higher: online enrolment.
Liu 2025	79	92 app arms	Post-test attrition 18.6%; follow-up 28.4%	Prediction intervals 2.4–68.3 and 3.0–83.7	Wide prediction intervals	Uptake 92.4%; adherence 61.8%. Lower attrition: reminders, human contact, <i>no</i> gamification. App vs wait-list RR 1.49.
Linardon 2023	70	9,258 allocated to a mindfulness app	Attrition 24.7%	—	Higher in large trials (38.7% when $n \geq 100$ per arm)	Same construct: assessment completion, not app-open.

Weights are the published meta-analytic weights inside those papers. They are not reproduced as a new forest plot of g , because the estimand is not Hedges g of a day-15 open, and because this review did not re-extract the primary cells.

Read Table 2 slowly. A trial that “retains” 74% to 81% of participants is retaining them for a form. The same participant may have stopped opening the programme on day 6. Wu 2021 found something that should have ended a thousand pitch decks: more persuasive-design features were associated with *larger* symptom effects and with *worse* study completion ($\beta = -0.0293$ on completion). Features that make a screenshot look busy are not the same as features that make a person come back on a working Tuesday.

Mohr 2019 is the trial that looks, at first glance, like a refutation of the cliff. In a modular suite with an eight-week window, median last use was day 56 of 56. After the official programme ended, 84% were still using it, at a median of 92 further days. Coaching increased downloads. Weekly recommendations increased sessions. Two facts keep this result in its box. The sample was 301 trial volunteers, not a store panel. Every arm was already inside a research protocol. High last-use in a factorial trial is evidence that a selected group, with scaffolding, can stay. It is not evidence that the 3.9% floor was a rounding error.

6.4 Moderator table

The five pre-specified moderators were reminders, human support, visible progress or streaks, messaging versus app, and language localisation. None of them has a new pooled Hedges g from this review, because the day-15 endpoint was not available in five comparable studies. What the published trial pools and the real-world contrasts do support is written below as a narrative moderator table.

Table 3. Pre-specified moderators

Moderator	Real-world signal	Trial-dropout signal	Pooled difference for D15/D30 same-day open	Verdict
Reminders	Not isolated in Baumei. Vendor claims of a three-fold lift among reminder-setters are UNVERIFIED.	Linardon 2020: lower attrition when participants were reminded. Liu 2025: lower attrition in trials that offered reminders.	Not estimable	Associated with lower <i>trial</i> dropout. Not shown to lift store-app day-15 open rates.
Human support	Employer platform with sessions: D30 53.3% vs Baumei 3.3%. AbleTo structured course: D30 45.1%. Pratap: clinician referral +40 days.	Torous 2020: dropout 11.7% with human feedback vs 34% without. Liu 2025: human contact, lower attrition. Mohr 2019: coaching increased downloads.	Not estimable as g	Strongest consistent signal. Moves real-world programmes into a different band.
Visible progress / streaks / gamification	Baumei: trackers 6.1% D30 vs breathing 0.0%. Peer support (a social form of visible progress) 8.9%, $n=2$.	Liu 2025: attrition <i>lower</i> when there was <i>no</i> gamification. Wu 2021: more persuasive-design features, worse completion.	Not estimable; requested contrast not supported	Mixed, and on the trial side currently pointed the wrong way. Do not claim that a streak counter raises day-15 retention.
Messaging vs app	Peer-support apps, $n=2$, best Baumei D30. PracticePal (messaging beside a human) high engagement in a 30-person rural pilot.	No head-to-head D15/D30 pool.	Not estimable	Thin. Promising as a <i>channel for human touch</i> , not as a substitute for it.
Language localisation	Mehrotra 2025: 2.9% of 350 apps available to Indian users offered an Indian language other than English.	PracticePal Hindi co-design; MITHRA audio/video for low literacy — both high completion inside a human wrap. No quantitative contrast.	Not estimable	Untested as a moderator of day-15 open rates. The missing cell is the finding.

The pre-specified question — the retention difference for programmes with visible progress counters — therefore gets this answer, and it is a null. There is no verified pooled difference in day-15 or day-30 same-day open rates for programmes that display a progress counter or a streak compared with programmes that do not. Trackers beat breathing in one panel. Gamification sat with worse trial retention in the largest modern trial synthesis. Those two facts do not add up to a product requirement for a flame icon.

6.5 Heterogeneity

Heterogeneity on the trial side is not a footnote. Torous 2020 reported $I^2 = 93.4\%$ for dropout. Liu 2025 reported prediction intervals that run from a few percent to more than two-thirds. That width means a new trial, drawn from the same universe, could land almost anywhere. Linardon 2023 found attrition of 24.7% overall and 38.7% once the sample per arm reached 100. Small, pretty trials keep their people. Large trials lose them. That pattern is a small-study effect by another name.

On the real-world side the heterogeneity is definitional. Same-day open, rolling use, module completion, last-use day, and questionnaire completion are five different y-axes. Pooling them would produce a precise number about nothing.

6.6 Publication bias and small-study effects

Torous 2020 ran Egger's test ($t = 3.9$, $p < 0.001$) and a bias adjustment that moved dropout from 26.2% to 47.8%. That is not a rounding error. It is a warning that the published trial record is kinder than the missing trial record. Liu 2025 did not erase that warning; it widened the prediction interval instead. This review did not run a new Egger test on day-15 open rates, because there was no new pool to test.

6.7 Risk of bias

Randomised trials in this literature typically have low risk on randomisation and high or some concerns on missing outcome data – which is circular, because missing outcome data *is* the dropout being measured. Detection of a questionnaire can be blinded. Detection of an app-open cannot be faked by the participant, but it can be defined in six incompatible ways by the analyst. The Baumel panel is at low risk of developer-wish bias and at high risk of Android-only and English-store bias. Employer and commercial cohorts are at high risk of selection: the denominator is already a person who works for a company that bought the programme, or who paid. Indian pilots (Bhat 2025; PracticePal 2025) are at low risk of cultural mismatch and at high risk of small-sample optimism.

RoB 2 judgements are not listed trial-by-trial here because this paper did not re-extract the 79 primary trials inside Liu 2025. Readers who need arm-level RoB should read that paper. The honest summary for this review: bias is not a reason to discard the cliff. Bias is a reason to believe the trial-dropout numbers are a floor on loss, not a ceiling.

6.8 GRADE

Table 4. GRADE by construct

Question	Estimate	Certainty	Reasons
What is the day-15 / day-30 same-day open rate for unguided store wellbeing programmes?	Median 3.9% / 3.3% (Baumel 2019)	Low	Single independent panel; Android-heavy; no 2019–2026 replication; indirectness for India and for messaging channels. Consistent with later rolling-retention and abandonment reviews.
What is assessment-dropout in trials of wellbeing / depression / anxiety apps?	19–26% post-test; ~28–36% later; bias-adjusted estimates higher	Moderate	Three independent meta-analyses, large k , consistent direction. Downgraded for inconsistency (I^2 and wide prediction intervals) and for indirectness to real-world opening.
Do reminders reduce loss?	Yes, for trial dropout. Unknown for store open rates.	Low for trial dropout as a causal claim; very low for store retention	Moderator analyses inside reviews, not a dedicated reminder RCT on day-15 open rates.
Does human support reduce loss?	Yes, across trial dropout and real-world guided cohorts.	Moderate for a directional claim; low for a precise increment	Convergent evidence from Torous, Liu, Mohr, Maistrello, Forman-Hoffman, Pratap, and the Indian human-wrap pilots. Residual confounding in observational workplace data.
Do visible progress counters raise day-15 or day-30 open rates?	No verified pooled lift. Trial-side gamification signal currently negative.	Very low for a benefit claim	Inconsistent metrics; possible harm signal in Liu 2025 and Wu 2021; Baumel tracker contrast is not a streak-counter experiment.
Does language localisation raise retention for Indian adults?	Unknown	—	No eligible contrast.

Certainty is low where the construct is the one this paper was asked to pool. That is not a failure of the search. It is a failure of the field to report the same endpoint twice.

07

Discussion

PULL QUOTE

Workplace programmes with a person in the loop report day-30 retention above 45%.

7.1 What the number means

Three point nine percent is not a rounding error and it is not a morality tale about lazy users. It is what an unguided store programme looks like when nobody is watching the participant and nobody is being paid to complete a form. The number has survived seven years without a second independent panel to replace it. That is a long time in a market that relaunches itself every quarter.

The number is also not a law of nature. An employer programme with a human session on the menu keeps about half its people to day 30. A structured commercial course keeps about 45%. A 30-person Hindi-language pilot with a counsellor in the next chair keeps almost everyone for the human programme and more than half the bot content. The cliff is steepest where the programme is a stranger, in a language that is not the user's first, with no one to notice that day 4 was skipped.

Trial dropout of one in four, or one in two after a bias correction, is a different sentence. It says: even when we recruit people who have agreed to be studied, we lose a quarter of the outcome data, and we probably lose more than we admit. It does not say that 75% of those people opened the programme on day 15. Wu's negative coefficient on completion is the machine-learning version of the same warning. Adding features can improve the loss you care about in the paper (symptom *g*) and worsen the loss you care about in production (people leaving). That is not a paradox. It is a regularisation problem. The model is fitting the screenshot.

Visible progress deserves a plain paragraph because a pooled difference was pre-specified and the evidence refused to provide one. A tracker, in Baumel, beat a breathing exercise. That may be because a tracker is a two-second rep with a mark on a page: small, dated, repeatable. A streak counter is a different object. It turns a miss into a catastrophe. A team lead in Noida misses one evening's rep because a release ran late. Next morning the flame reads zero. Now the app feels like a report card, so she leaves it shut. Liu 2025 found that trials without gamification had lower attrition. The honest product reading is not "kill all progress". It is "show the work, do not ransom the relationship to an unbroken flame".

Reminders help in trials. They may help in the wild. They have not been isolated as a cause of a higher day-15 open rate in a store panel. A reminder is a prompt. It can start a rep; it cannot do the rep. In entrepreneurial terms it is a re-activation campaign paid for with the user's attention. Used like a fire alarm — rare, specific, tied to a next rep — it is a design feature: "Two minutes after lunch: write one good thing from this morning." Used like a needy friend — the fourth "We miss you!" of the week — it is how an app gets muted.

Human support is the feature that keeps surviving contact with every dataset. Torous: 11.7% dropout with feedback versus 34% without. Liu: human contact in the lower-attrition set. Pratap: a clinician's referral worth 40 extra days. Maistrello: a workplace channel with sessions, day-30 at 53%. The Indian pilots wrap a digital tool in a group or a counsellor and then report completion figures that look like a different industry. The picture here is a desk, not a dashboard. On Wednesday a colleague in Hyderabad notices that the analyst beside her skipped yesterday's rep. Over chai she asks, "Want to do today's together?" They finish it in two minutes before stand-up. That is social support at work: one person who notices a miss and makes the next rep easy. A programme that cannot survive a missed Tuesday needs a person, not another badge.

7.2 Limitations

This review did not dual-screen every raw database hit. The PRISMA counts are reconstructed. A registered update with two independent reviewers should replace that reconstruction.

The day-15 endpoint is almost unpublished outside one panel. Narrative synthesis of a single panel plus heterogeneous cousins is not a meta-analysis, and this paper has not pretended otherwise.

Trial-dropout pools overlap. Some trials sit in Torous and in Linardon and in Liu. Quoting all three is useful triangulation. Adding their k as if they were independent would be double-counting.

Android panels miss iOS. Store panels miss WhatsApp. English-language literature misses the languages in which most Indian professionals actually scold themselves. Employer cohorts miss the informal workforce. University trials miss the 6 a.m. shift.

Unverified commercial dashboards were excluded. That choice keeps the numbers honest and throws away whatever signal lives in those dashboards. Both sides of that trade should be named.

The protocol planned Hedges g , a prediction interval, Egger and trim-and-fill on a new pool. Those statistics were not computed because the pool was not eligible. Reporting them on a homemade average of Baumel, Kozlov and Maistrello would have been a more precise way to be wrong.

7.3 What this means for an Indian workplace programme

Take an HR head in Pune who buys an unguided, store-style app for the whole floor. She should expect the 4% floor, not the trial's 75%. By the second week the icon sits in a folder nobody opens. Now run the same launch with a budget for a human hand-off. For the first ten days a team lead, not a flame icon, asks after stand-up: "Done today's rep?" The reminder arrives on the message thread people already open, in the language they actually think in, and names tomorrow's one rep: "Tomorrow, before your first call, name one feeling in one word." She does not buy gamification and call it engagement. She counts daily reps. She publishes her own day-15 and day-30 same-day open rates, in Hindi and in English, with a denominator and a date. Until someone does that, every India slide that quotes 3.9% is quoting a Californian Android panel and hoping the local train is similar. It is not.

7.4 Research gaps, naming the India gap

The field still does not report one endpoint. Day-15 same-day open, rolling month-1 use, module completion, last-use day, and assessment completion travel under the single passport of "engagement". They should be stamped separately.

There is no second independent multi-app panel. Seven years is long enough. A 2026 panel that includes iOS, messaging programmes, and at least the ten most-used Indian languages would settle whether 3.9% was a moment or a climate.

Moderator trials that cross reminders, human dose, and progress displays as separate factors, with day-15 open as a pre-specified outcome, do not exist at adequate scale. Mohr 2019 is the closest ancestor and did not report that endpoint.

The India gap, named as a gap. No published study reports a day-15 or day-30 same-day open rate for an adult consumer wellbeing programme delivered in an Indian language to working-age users. Ghosh 2023 came closest on the unguided side and found full completion of 5.5% — a trial number that looks like the store floor, not like Torous. Bhat 2025 and PracticePal 2025 show what a human wrap can do in a group or a village clinic, and they still do not give a commuter in Gurugram a day-15 curve. Mehrotra 2025 shows that the shelf facing that commuter is 97% not in their other language and 89% without a cited study. CTRI holds registrations for community digital-care programmes. It does not hold a consumer retention curve. That is the gap this paper was asked to name, and it remains open on 22 September 2026.

08

FAQ

Q1 Do most wellbeing apps really lose 96% of people by day 15?

Yes, on the only independent multi-app measure of same-day opens: median day-15 retention 3.9% across 93 programmes. That is a store-app number, not a workplace-coach number. Guided programmes sit nearer 45–53% at day 30.

Q2 If a trial says 75% were “retained”, did they open the app for a month?

No. That figure is usually the share who completed a questionnaire. Torous pooled 26.2% dropout across 18 trials; Liu pooled 18.6% post-test attrition across 79. Completing a form is not an open on day 15.

Q3 Will a streak counter keep our employees on the programme?

Not on present evidence. Trackers beat breathing exercises in one panel (6.1% versus 0.0% at day 30). The largest recent trial synthesis found *higher* attrition where gamification was present. Show the work. Do not ransom Monday to an unbroken flame.

Q4 What actually bends the curve?

A human. Dropout was 11.7% with human feedback and 34% without in Torous. Employer programmes with sessions report day-30 retention above 45%. Reminders help in trials. Language that matches the user’s first tongue is untested as a number and obvious as a courtesy.

Q5 Is there an Indian day-15 figure we can put in a board slide?

No published one. Treat 3.9% as the unguided import default, 45% as the guided-workplace default, and your own dashboard — once you publish it with a denominator — as the only local truth.

09

References

- Baumel A, Muench F, Edan S, Kane JM. Objective user engagement with mental health apps: systematic search and panel-based usage analysis. *J Med Internet Res*. 2019;21(9):e14567. <https://doi.org/10.2196/14567>
- Torous J, Lipschitz J, Ng M, Firth J. Dropout rates in clinical trials of smartphone apps for depressive symptoms: a systematic review and meta-analysis. *J Affect Disord*. 2020;263:413-419. <https://doi.org/10.1016/j.jad.2019.11.167>
- Linardon J, Fuller-Tyszkiewicz M. Attrition and adherence in smartphone-delivered interventions for mental health problems: a systematic and meta-analytic review. *J Consult Clin Psychol*. 2020;88(1):1-13. <https://doi.org/10.1037/ccp0000459>
- Liu C, Torous J, Fuller-Tyszkiewicz M, Messer M, Anderson C, Soliman OM, Linardon J. Uptake, adherence, and attrition in clinical trials of depression and anxiety apps: a systematic review and meta-analysis. *JAMA Psychiatry*. 2026;83(1):43-50. Published online 19 November 2025. <https://doi.org/10.1001/jamapsychiatry.2025.3439>
- Linardon J. Rates of attrition and engagement in randomized controlled trials of mindfulness apps: systematic review and meta-analysis. *Behav Res Ther*. 2023;170:104421. <https://doi.org/10.1016/j.brat.2023.104421>
- Wu A, Scult MA, Barnes ED, Betancourt JA, Falk A, Gunning FM. Smartphone apps for depression and anxiety: a systematic review and meta-analysis of techniques to increase engagement. *npj Digit Med*. 2021;4:20. <https://doi.org/10.1038/s41746-021-00386-8>
- Pratap A, Neto EC, Snyder P, et al. Indicators of retention in remote digital health studies: a cross-study evaluation of 100,000 participants. *npj Digit Med*. 2020;3:21. <https://doi.org/10.1038/s41746-020-0224-8>
- Kozlov E, Bantum E, Pagano I, et al. The reach, use, and impact of a free mHealth mindfulness app in the general population: mobile data analysis. *JMIR Ment Health*. 2020;7(11):e23377. <https://doi.org/10.2196/23377>
- Mohr DC, Schueller SM, Tomasino KN, et al. Comparison of the effects of coaching and receipt of app recommendations on depression, anxiety, and engagement in the IntelliCare platform: randomized controlled trial. *J Med Internet Res*. 2019;21(8):e13609. <https://doi.org/10.2196/13609>
- Anton MT, Mochari-Greenberger H, Andreopoulos E, Pande RL. A late-binding, instant-access service for digitally supported behavioral health: retrospective cohort of engagement and retention. *JMIR Form Res*. 2021;5(9):e27570. <https://doi.org/10.2196/27570>
- Maistrello G, Phillips W, Lee DS, Romanelli RJ. Utilization and engagement with an employer-provided digital mental health platform: an international perspective. *Digit Health*. 2024. <https://doi.org/10.1177/20552076241277180>
- Kidman PG, Curtis RG, Watson A, Maher CA. When and why adults abandon lifestyle behaviour and mental health mobile apps: scoping review. *J Med Internet Res*. 2024;26:e56897. <https://doi.org/10.2196/56897>
- Fleming T, Bavin L, Lucassen M, Stasiak K, Hopkins S, Merry S. Beyond the trial: systematic review of real-world uptake and engagement with digital self-help interventions for depression, low mood, or anxiety. *J Med Internet Res*. 2018;20(6):e199. <https://doi.org/10.2196/jmir.9275>
- Borghouts J, Eikey E, Mark G, et al. Barriers to and facilitators of user engagement with digital mental health interventions: systematic review. *J Med Internet Res*. 2021;23(3):e24387. <https://doi.org/10.2196/24387>
- Boucher EM, Raiker JS. Engagement and retention in digital mental health interventions: a narrative review. *BMC Digit Health*. 2024;2:52. <https://doi.org/10.1186/s44247-024-00105-9>
- Ghosh A, Cherian RJ, Wagle S, et al. An unguided, computerized cognitive behavioral therapy intervention (TreadWill) in a general population of Indian users: randomized controlled trial. *J Med Internet Res*. 2023;25:e41005. <https://doi.org/10.2196/41005>
- Bhat A, et al. Multiuser application for the diagnosis and treatment of depression in women's self-help groups: pilot randomized controlled trial. *JMIR Form Res*. 2025;9:e68052. <https://doi.org/10.2196/68052>
- A conversational agent (PracticePal) to support the delivery of a brief behavioral activation treatment for depression in rural India: development and pilot-testing study. *JMIR Form Res*. 2025;9:e73563. <https://doi.org/10.2196/73563>
- Mehrotra S, Tripathi R, Sengupta P, et al. Evaluating characteristics and quality of mental health apps available in app stores for Indian users: systematic app search and review. *JMIR mHealth uHealth*. 2025;13:e79238. <https://doi.org/10.2196/79238>
- Nwosu A, Boardman S, Husain MM, Doraiswamy PM. Digital therapeutics for mental health: is attrition the Achilles heel? *Front Psychiatry*. 2022;13:900615. <https://doi.org/10.3389/fpsyt.2022.900615>
- Flett JAM, Hayne H, Riordan BC, Thompson LM, Conner TS. Mobile mindfulness meditation: a randomised controlled trial of the effect of two popular apps on mental health. *Mindfulness*. 2019;10:863-876. <https://doi.org/10.1007/s12671-018-1050-9>
- Zhang Y, et al. Long-term participant retention and engagement patterns in an app and wearable-based multinational remote digital depression study. *npj Digit Med*. 2023;6:25. <https://doi.org/10.1038/s41746-023-00749-3>
- Leo AJ, et al. A digital mental health intervention in chronic pain: ancillary survival analysis of user retention. *JMIR*. 2022. <https://doi.org/10.2196/38284>
- Hallenbeck HW, et al. PTSD Coach version 3.1: a comparison of engagement. *JMIR mHealth uHealth*. 2022. <https://doi.org/10.2196/34744>

Owen JE, Jaworski BK, Kuhn E, Makin-Byrd KN, Ramsey KM, Hoffman JE. mHealth in the wild: using novel data to examine the reach, use, and impact of PTSD Coach. *JMIR Ment Health*. 2015;2(1):e8. <https://doi.org/10.2196/mental.3935>

Linardon J, Torous J, Firth J, Cuijpers P, Messer M, Fuller-Tyszkiewicz M. Current evidence on the efficacy of mental health smartphone apps for symptoms of depression and anxiety: a meta-analysis of 176 randomized controlled trials. *World Psychiatry*. 2024;23(1). <https://doi.org/10.1002/wps.21183>

About the author

ABOUT THE AUTHOR

Dr Atul Aswani; Mumbai Psychiatrist and trained as a Results Coach

Atul Aswani is a Mumbai Psychiatrist and trained as a Results Coach. He builds EmotionApp, a non-clinical emotional-fitness programme for Indian professionals: countable daily reps, an AI coach with human hand-off, WhatsApp-first delivery, twenty-three Indian languages, India-resident data, DPDP-compliant, under ISO 9001 and ISO 13485. He works from Mumbai. He is interested in practices people still open on a working Tuesday.

HOW TO CITE

Aswani A. The Attrition Cliff: Retention in digital wellbeing interventions: a meta-analysis of day-15 and day-30 retention and its design moderators, 2019–2026. EmotionApp Research Paper 2. Mumbai: Tranquilo Meditek Sytems Private Limited; 2026. <https://emotionapp.in/research/attrition-cliff-retention-digital-wellbeing>

LAST UPDATED

Last updated 22 September 2026

NOTE

EmotionApp is a non-clinical emotional-fitness coaching platform. This paper describes practices, techniques and coaching. It is not therapy, treatment, diagnosis or medical advice.

© 2026 Tranquilo Meditek Sytems Private Limited, Mumbai · emotionapp.in

SAFETY NOTE

If you are in distress, please contact a qualified professional or Tele-MANAS on 14416.