Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

App-based symptom monitoring in dermatology: pooled prospective trials to quantify engagement, retention and ecological momentary assessment trajectories (2018–2025)

Tassilo Dege
University Hospital Würzburg
Igor Bibi
University Medical Centre Mannheim
Tizian Claus Dege
TU Darmstadt
Jiahong Yuan
University Hospital Würzburg
Caroline Glatzel
University Hospital Würzburg
Jan-Hendrik Maiwald
University Hospital Würzburg
Matthias Goebeler
University Hospital Würzburg
Patrick-Pascal Struntz
University Hospital Würzburg
Victor Olsavszky
University Medical Centre Mannheim
Christian Hametner
University Hospital Würzburg
Astrid Schmieder
Schmieder_a@ukw.de
University Hospital Würzburg

Article

Keywords: EMA, engagement, retention, psoriasis, eczema, digital endpoints, data harmonization

Posted Date: March 8th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8888185/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: Competing interest reported. TD: Travel grants or speaker fees: Leo Pharma, Recordati Rare Diseases, Johnson & Johnson, Sanofi. IB: None declared. TCD: None declared. JY: None declared. CG: None declared. JHM: None declared. MG: Scientific advisory board/ speakers bureau: Almirall, Argenx, Biotest, Fresenius, GSK, Janssen, Leo Pharma, Lilly, Novartis, UCB. PPS: Research support: Novartis, Abbvie, and Chugai. Travel grants or speaker fees: Abbvie, UCB, Janssen, Novartis VO: Scientific advisory board/ speakers bureau: BMS, Johnson & Johnson, and UCB. CH: None declared. AS: research support/ clinical trials: Abbvie, Boehringer-Ingelheim, Celgene, Eli Lilly, Janssen-Cilag, LEO Pharma, Merck, Novartis, Pfizer; Scientific advisory board/ speakers bureau: Abbvie, Almirall, Hermal, Janssen, LEO, and Novartis, UCB.

Abstract

Background: Digital health monitoring delivered via smartphone applications and ecological momentary assessment (EMA) can capture high-frequency symptom trajectories in chronic inflammatory skin diseases, but real-world engagement and retention vary substantially across studies and can complicate inference. We pooled five prospective clinical trials to harmonize clinic visits, patient-reported outcomes (PROs), and smartphone app–derived EMA logs and to quantify engagement, retention, and clinical anchoring of EMA signals.

Methods: We harmonized five prospective dermatology studies (psoriasis and chronic hand/foot eczema spanning from 2018 to 2025, including two randomized controlled trials and three non-randomized studies, conducted at University Hospital Würzburg and University Medical Center Mannheim, into a relational dataset with patient-, visit-, and EMA-level tables. All participants were provided access to the smartphone monitoring app; actual use was voluntary and quantified from timestamped logs. In the app-assigned cohort, baseline characteristics were summarized overall and by engagement group (no use vs any use). Early engagement was quantified as active days in the first 28 days after first activity and retention as time from first to last observed activity, summarized with Kaplan–Meier curves. EMA pruritus and pain trajectories were summarized weekly over 26 weeks and anchored to clinic symptom assessments at baseline and 6 months. Associations between early engagement and 6-month improvement were evaluated using adjusted models.

Results: The harmonized baseline clinical dataset included 550 app-assigned participants (95.6% psoriasis). Baseline disease burden was moderate (median Dermatology Life Quality Index (DLQI) 6 [IQR (Interquartile range) 2–14]; median pruritus 2 [1–5]; median pain 1 [0–3]). Baseline characteristics were broadly comparable between engagement groups (no use n=211 vs any use n=339) with similar baseline PROs and symptoms. Engagement in the first 28 days was low and right-skewed (median 2 active days [1–4]), and retention showed early drop-off with heterogeneous retention across trials (retained 68.4% at day 28; 29.8% at day 182). Weekly EMA trajectories aligned with clinic anchors and separated by clinical response strata (ΔDLQI quintiles). Higher early engagement showed a graded association with greater adjusted 6-month improvement, with wider uncertainty at very high engagement.

Conclusions: In a multi-trial pooled dermatology cohort, EMA symptom tracking was feasible and clinically interpretable when anchored to visit-based assessments, while engagement and retention varied substantially across studies. Harmonized, reproducible data structures integrating clinic-based assessments, PROs, and EMA logs can support effective trialing of digital interventions and enable robust quantification of engagement heterogeneity and between-visit symptom trajectories.

Introduction

Chronic inflammatory skin diseases such as psoriasis and chronic hand/foot eczema are both skin conditions that impose substantial symptom burden and quality-of-life impairment, often characterized

by fluctuating pruritus, pain, and psychosocial distress.[1] Traditional clinical follow-up schedules capture outcomes sparsely, limiting the ability to characterize short-term variability, early changes, and between-visit symptom dynamics.[2] Digital interventions delivered via smartphones can extend monitoring beyond clinic visits and enable integration of patient-reported outcomes (PROs) with high-frequency symptom tracking through ecological momentary assessment (EMA).

However, digital intervention trials pose unique challenges for effective trialing in real-world settings.[3] Engagement and retention commonly show early drop-off and differ across study contexts, affecting statistical power, representativeness, and interpretability of outcomes.[4-7] Moreover, pooling evidence across studies is often hindered by heterogeneity in variable definitions, coding, and data structures.[8] Harmonization across studies can strengthen inference and enable comparative evaluation across trial types (randomized controlled trial (RCT) and non-RCT) and disease populations.[9]

In this study, we pooled and harmonized five prospective dermatology studies conducted at two academic centers to (i) establish a reproducible integrated dataset structure linking clinic visits, PROs, and EMA logs; (ii) quantify engagement and retention heterogeneity across trials; (iii) evaluate whether EMA symptom trajectories are clinically interpretable when anchored to visit-based assessments; and (iv) assess associations between early engagement intensity and 6-month clinical improvement. Our goal is to provide a framework and quantitative evidence supporting effective trialing of digital interventions using harmonized, multi-source datasets in dermatology. Beyond dermatology, the relational data model and engagement metrics offer a reusable template for harmonizing app-derived logs with clinic-based and PRO data across heterogeneous digital health studies.

Methods

Study design and data sources

This is a pooled, harmonized analysis of five prospective dermatology studies (psoriasis and chronic hand/foot eczema; RCT and non-RCT) conducted at (1) the Department of Dermatology, Venereology and Allergology, University Hospital Würzburg, Germany and (2) the Department of Dermatology, Venereology and Allergology, University Medical Center Mannheim, Germany. Each study included clinic-based assessments and smartphone-based symptom tracking using study-specific app versions and a clinician dashboard (Supplementary Figure S1). Participants provided written informed consent before enrollment. Ethics approvals were obtained from the Ethics Committee of the University of Würzburg (98/23_skpf; 129/22) and the Medical Faculty Mannheim, Heidelberg University (2022-502; 2021-895; 2020-515N-MA; 2017-655N-MA).

Harmonization workflow and pooled dataset structure

Raw clinical and digital log data from the contributing studies (Figure 1A) were ingested and harmonized through a standardized pipeline (Figure 1B) comprising (i) raw data ingest, (ii) cleaning and quality

control, and (iii) variable mapping to a shared data dictionary. The final pooled dataset followed a reproducible structure with a patient-level table and long-format tables for clinical visits and EMA entries (Figure 1C). Harmonization procedures included standardization of variable names, coding reconciliation (e.g., categorical encodings), and consistency checks across timepoints.

Participants and analysis cohorts

Participants assigned to app follow-up were provided access to the study smartphone application at baseline and were instructed on its use for symptom monitoring (by randomization in 2 trials and by protocol assignment in 3 non-RCTs). Use was voluntary and quantified from timestamped app logs. For the present analysis, the app-assigned cohort comprised participants flagged as assigned to app follow-up.

Measures

Clinical and PRO measures: Baseline variables included age, sex, diagnosis (psoriasis or eczema), Body Mass Index (BMI), smoking status, obesity, depression and hypertension (binary), baseline Dermatology Life Quality Index (DLQI; 0–30, higher = worse quality-of-life impairment)[10,11], Hospital Anxiety and Depression Scale (HADS; anxiety and depression subscales, each 0–21, higher = worse symptoms) – anxiety (A), HADS-depression (D),[12] and symptom numerical rating scales (NRS) for pruritus and pain. Follow-up outcomes included 6-month DLQI and derived change scores (ΔDLQI = baseline − 6 months; positive values indicate improvement) used for response stratification and modeling.

Engagement and retention

- Engagement group: Participants were classified as “No use” vs “Any use” based on presence of at least one logged activity in the app.
- Early engagement intensity: active days with ≥1 entry in the first 28 days after first activity.
- Retention: time from first to last observed activity day (days), summarized using Kaplan–Meier curves overall and by trial
- Usage timing: heatmap of unique active users by weekday and hour of day (Europe/Berlin time zone).
- Session structure and revisit frequency: Median per-session feature counts (home/history/chat opens; app pauses) and inter-session intervals (days between consecutive session-days per user; capped at 30 for visualization).
- EMA symptom trajectories: EMA pruritus and pain entries were summarized weekly (weeks 0–26) and compared with clinic anchors at baseline and 6 months. EMA trajectories were stratified by clinical response quintiles based on ΔDLQI.

Statistical analysis

Baseline characteristics were summarized overall (Table 1) and by engagement group (Table 2), using mean (SD) or median [IQR] for continuous variables and n (%) for categorical variables. Unadjusted baseline group comparisons used Mann–Whitney U tests for continuous variables and chi-square or Fisher's exact tests for categorical variables (Table 2). Retention was summarized using Kaplan–Meier curves (Figure 2B).

Associations between early engagement (active days in the first 28 days) and adjusted 6-month improvement were evaluated using regression models adjusted for prespecified covariates (e.g., age, baseline DLQI, and trial indicators), visualized as a dose–response curve with uncertainty bands (Supplementary Figure S3). Analyses were conducted in Python using pandas, numpy, scipy, and lifelines/statsmodels; code and harmonization logic were maintained in a reproducible analysis pipeline.

Missing data

Analyses used available data per variable and timepoint; denominators are reported in tables/figure panels. Because the pooled setting integrates multiple studies and data capture modalities, missingness and coding heterogeneity were expected and addressed through harmonization, and adjustment for trial indicators in modeling where applicable.

Results

Cohort and pooled dataset

Five prospective dermatology studies (psoriasis and chronic hand/foot eczema; RCT and non-RCT) contributed to a harmonized pooled dataset linking clinic visits, PROs, and EMA logs (Figure 1)[13-16]. The harmonized baseline clinical dataset comprised 550 app-assigned participants with baseline visit data (Table 1). Most participants had psoriasis (526/550; 95.6%), with a smaller eczema subgroup (24/550; 4.4%). Participants had a mean age of 47 years (SD 14) and were predominantly male (315/550; 57.3%), with sex missing/undeclared in 11/550 (2.0%). Baseline burden was moderate (median DLQI 6 [IQR 2–14]; pruritus NRS 2 [1–5]; pain NRS 1 [0–3]; HADS-A 6 [3–10]; HADS-D 4 [2–7]). Comorbidity burden was notable (obesity 214/550; 38.9%; hypertension 162/550; 31.6%), and 194/550 (35.3%) had a documented treatment change during follow-up (baseline–6 months).

Table 1: Baseline characteristics of the assigned clinical cohort.

Baseline characteristic	Overall (n=550)
Female	224 (40.7%)
Male	315 (57.3%)
Sex missing/undeclared	11 (2.0%)
Age, years (mean (SD))	47 (14)
BMI, kg/m² (median [IQR])	28.1 [24.8–32.7]
Smoker	217 (41.0%)
Obesity	214 (38.9%)
Depression	71 (13.9%)
Hypertension	162 (31.6%)
DLQI, median [IQR]	6 [2–14]
HADS-A, median [IQR]	6 [3–10]
HADS-D, median [IQR]	4 [2–7]
Pruritus NRS, median [IQR]	2 [1–5]
Pain NRS, median [IQR]	1 [0–3]
Treatment change during follow-up	194 (35.3%)

Baseline demographic and clinical characteristics for participants assigned to app follow-up. Continuous variables are reported as mean (SD) or median [IQR] as appropriate; categorical variables are reported as n (%). Abbreviations: BMI, body mass index; DLQI, Dermatology Life Quality Index; HADS, Hospital Anxiety and Depression Scale; NRS, numerical rating scale; IQR, interquartile range.

Baseline comparability by app engagement group

Within the app-assigned cohort, participants were stratified into “No use” (n=211) and “Any use” (n=339), defined as having 0 versus ≥1 timestamped logged app activity (Table 2). Baseline characteristics were largely similar between groups. Median age was 50 [34–59] in “No use” and 48 [35–58] in “Any use” (p=0.588). Baseline PROs and symptoms were comparable, including DLQI (median 4 [1–14] vs 6 [2–14], p=0.078), HADS-A (p=0.986), HADS-D (p=0.715), pruritus NRS (p=0.443), and pain NRS (p=0.606). Two differences were observed: obesity was more prevalent in “No use” (45.0% vs 35.1%, p=0.024), and treatment change during follow-up (baseline–6 months) occurred more frequently in “No use” (43.1% vs 30.4%, p=0.003). Sex distribution was similar among participants with recorded sex; however, sex was missing/undeclared more often in “No use” (4.7% vs 0.3%).

Table 2: Baseline characteristics by engagement group.

Characteristic	No use (n=211)	Any use (n=339)	p value
Sex (recorded), n (%)			1.0
Female	84 (41.8%) (n=201)	140 (41.4%) (n=338)	
Male	117 (58.2%) (n=201)	198 (58.6%) (n=338)	
Sex missing/undeclared, n (%)	10 (4.7%)	1 (0.3%)	
Age, years (median [IQR])	50 [34–59]	48 [35–58]	0.588
BMI, kg/m² (median [IQR])	29.0 [25.5–33.2]	27.5 [24.5–32.5]	0.098
Smoker, n (%)	83 (42.6%) (n=195)	134 (40.1%) (n=334)	0.584
Obesity, n (%)	95 (45.0%) (n=211)	119 (35.1%) (n=339)	0.024
Depression, n (%)	27 (13.7%) (n=197)	44 (14.0%) (n=315)	1.0
Hypertension, n (%)	72 (36.5%) (n=197)	90 (28.6%) (n=315)	0.064
DLQI, median [IQR]	4 [1–14]	6 [2–14]	0.078
HADS-A, median [IQR]	6 [3–9]	6 [3–10]	0.986
HADS-D, median [IQR]	4 [2–8]	4 [2–7]	0.715
Pruritus (NRS), median [IQR]	2 [1–5]	2 [1–5]	0.443
Pain (NRS), median [IQR]	1 [0–3]	1 [0–4]	0.606
Treatment change during follow-up	91 (43.1%) (n=211)	103 (30.4%) (n=339)	0.003

Baseline characteristics stratified by engagement group (“No use” vs “Any use”) within the app-assigned cohort. Continuous variables are reported as median [IQR]; categorical variables are reported as n (%). P values reflect unadjusted group comparisons (Mann–Whitney U test for continuous variables; chi-square or Fisher’s exact test for categorical variables). Sex is reported as n (%) of the full group denominator, with missing/undeclared sex shown separately. Variable-specific denominators are shown where applicable.

Engagement intensity, retention, and usage timing

Among users with log data, early engagement in the first 28 days was low on average and right-skewed across trials, with trial-specific user counts shown in Figure 2A (Trial 1 n=40; Trial 2 n=108; Trial 3 n=30; Trial 4 n=110; Trial 5 n=83). Across the pooled “Any use” group, early engagement corresponded to a median of 2 active days [IQR 1–4] in the first 28 days (computed from logs).

Retention curves indicated an early drop-off followed by more gradual decline among continuing users. Across users, retention probabilities were 68.4% at day 28, 50.1% at day 84, and 29.8% at day 182, with marked heterogeneity between trials (Figure 2B).

Logging activity exhibited clear diurnal and weekly patterns: usage was lowest overnight (0:00–6:00), increased from early morning, and clustered during daytime and early evening (roughly 8:00–20:00). Peaks were most pronounced around late morning and early evening hours. Across weekdays, activity was broadly distributed, with a marked increase on Sundays (particularly late morning to afternoon and early evening), whereas Saturdays appeared comparatively lower. These temporal regularities support real-world feasibility of EMA-style symptom tracking and indicate that usage is shaped by daily routines (Figure 2C).

Session structure and revisit frequency

Session-level analyses indicated predominantly task-focused app interactions. The home page was opened a median of 5 times per session, whereas history and chat pages were rarely accessed (median 0 each), and app pauses occurred infrequently (median 1) (Figure 3A). Revisit patterns were intermittent, with inter-session intervals clustering around approximately weekly use (median 8 days; Figure 3B), complementing retention analyses by describing how active users returned to the app over time.

Clinical response at 6 months across trials (DLQI anchor)

Six-month change in DLQI (ΔDLQI; baseline − 6 months, positive values indicate improvement) showed overall improvement with heterogeneity in response distributions across the contributing studies (Figure 4). Subgroup estimates were broadly consistent with improvement across disease and trial strata, and by therapy stability (stable therapy vs therapy change) (Supplementary Figure S2).

EMA anchoring and associations with clinical improvement

Weekly EMA symptom trajectories for pruritus and pain were summarized over 26 weeks and anchored to clinic-based symptom assessments at baseline and 6 months (Figure 5). In the left panels (overall cohort), weekly EMA means were largely stable over follow-up with modest week-to-week variability, and the clinic anchor points fell within the range of contemporaneous EMA levels, supporting clinical interpretability relative to visit-based assessments.

Separation between Q1 and Q5 was already visible within the first few weeks after first logging and persisted through week 26. The difference was more pronounced and consistently visible for pruritus, with higher EMA pruritus levels in Q1 and lower, generally declining levels in Q5. For pain, separation was also evident but less pronounced and more variable over time. Clinic anchor points at baseline and 6 months mirrored this ordering (higher in Q1, lower in Q5), supporting EMA-derived symptom trajectories as a sensitive between-visit endpoint capturing clinically meaningful differences in symptom burden over time.

In adjusted models, early engagement intensity (active days with ≥1 entry during the first 28 days after first use) was associated with change outcomes over follow-up (Supplementary Figure S3). The fitted curves indicate a graded pattern in which higher early engagement corresponded to greater visit-based DLQI improvement and more favorable EMA-derived symptom change proxies for itch, pain, and skin.

Confidence intervals widened substantially at the upper end of engagement, consistent with fewer participants exhibiting very frequent early use. These patterns reflect observational associations and should not be interpreted causally.

Discussion

In this pooled harmonized analysis of five prospective dermatology studies integrating clinic visits, PROs, and EMA logs, we demonstrate three main findings relevant to effective trialing of digital interventions. First, engagement in real-world app-based monitoring was highly right-skewed and retention showed early drop-off, with substantial heterogeneity across contributing trials. Second, despite engagement variability, weekly EMA symptom trajectories for pruritus and pain were clinically interpretable when anchored to visit-based symptom assessments, supporting EMA as a meaningful between-visit endpoint. Third, higher early engagement intensity was associated with greater adjusted 6-month improvement, showing a graded relationship with wider uncertainty at very high engagement levels; these observational comparisons should not be interpreted causally.

A major strength of this work is the harmonization and reproducible dataset structure spanning multiple trial types (RCT and non-RCT), disease groups, and centers. The standardized patient-level and long-format visit/EMA tables (Figure 1C) enabled consistent estimation of engagement metrics, retention, and symptom trajectories across studies. This addresses a key barrier in digital medicine research: lack of standardization across interventions and datasets, which limits comparability and replication[17]. The engagement and retention patterns (Figure 2) are consistent with a common digital health phenomenon —early discontinuation among a large fraction of users[18]—underscoring the importance of explicitly quantifying engagement distributions rather than relying on average use.

Our findings also support EMA as a clinically meaningful signal, consistent with previous research[19-21]. EMA trajectories aligned with clinic anchors and separated early after first logging between clinical response strata (Figure 5). Participants in the poorer-response stratum showed persistently higher EMA symptom levels, whereas those in the better-response stratum showed lower and, for pruritus, generally declining trajectories across follow-up. This early and sustained separation suggests that high-frequency symptom tracking captures between-visit differences in symptom burden that may be missed by sparse visit schedules. Such between-visit information could improve understanding of early response dynamics and may support more sensitive endpoint definitions, adaptive intervention strategies, or more efficient trial designs.

Several limitations should be considered. First, engagement is not randomized; therefore, associations between early engagement and outcomes should not be interpreted causally. Engagement intensity likely reflects a mixture of motivation, disease burden, treatment changes, and contextual factors. Second, pooling across trials introduces heterogeneity in recruitment, study procedures, and measurement frequency, which can influence retention and EMA coverage. We addressed this by harmonization and adjustment using trial indicators, but residual heterogeneity is possible. Third, the

eczema subgroup was small relative to psoriasis, limiting disease-specific inference. Finally, some variables differed in definition or category coding across trials and required harmonization, which may leave residual inconsistencies despite quality control.

Future work should evaluate causal mechanisms linking digital exposure to outcomes using designs such as micro-randomized trials, encouragement designs, or instrumental-variable approaches, and should further standardize digital endpoints and engagement definitions. Integrating additional digital phenotyping signals (e.g., passive sensing) and exploring personalized engagement strategies may improve retention and increase the utility of EMA-derived endpoints.

In summary, we harmonized five prospective dermatology studies across trial designs and centers into a reproducible relational dataset linking clinic-based assessments, PROs, and EMA logs, enabling comparable quantification of engagement, retention, and between-visit symptom trajectories. EMA symptom tracking was clinically interpretable when anchored to visit-based assessments, while engagement and retention were highly heterogeneous and should therefore be reported transparently using standardized, clearly defined metrics. This pooled framework provides a practical foundation for more standardized evaluation of digital endpoints and digital interventions in real-world settings.

Declarations

Acknowledgments

We sincerely thank all patients.

Author Contributions

TD: Conceptualization, Data curation, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing.

IB: Data curation, Investigation, Methodology, Validation, Writing – review & editing.

TCD: Data curation, Investigation, Methodology, Validation, Visualization, Writing – review & editing.

JY: Data curation, Investigation, Methodology, Visualization, Writing – original draft.

CG: Data curation, Investigation, Methodology, Validation.

JHM: Data curation, Investigation, Methodology, Validation.

MG: Supervision, Validation, Writing – review & editing.

PPS: Investigation, Validation, Writing – review & editing.

VO: Data curation, Investigation, Methodology, Validation.

CH: Investigation, Methodology, Validation, Writing – review & editing.

AS: Conceptualization, Supervision, Validation, Investigation Visualization, Writing – original draft, Writing – review & editing.

Funding

TD has been supported by the clinician scientist program of the Interdisciplinary Center of Clinical Research (IZKF; project number: Z-2/CSP-25), Medical Faculty, University of Würzburg. CG is clinician scientist supported by the TWINSIGHT Clinician Scientist Program (project number: TWINSIGHT-09) at the Medical Faculty, University of Würzburg, that is funded by the Else-Kröner-Fresenius Foundation. The study is funded by the German Federal Ministry of Education and Research (BMBF), consortium project HybridVita (grant number: 16SV8903). In addition, the study was supported by an unconditional grant from Novartis GmbH.

Medical Writing/Editorial Assistance

None.

Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Ethical Approval

The study was conducted in accordance with the principles of the Declaration of Helsinki and approved by the Ethics Committee of the University of Würzburg (reference number: (98/23_skpf **;** 129/22) and of the Medical Faculty Mannheim, Heidelberg University (reference number: 2022-502**;** 2021-895**;** 2020-515N-MA**;** 2017-655N-MA). All participants provided written informed consent prior to enrollment. Participants received no compensation for their participation.

Competing Interests

TD: Travel grants or speaker fees: Leo Pharma, Recordati Rare Diseases, Johnson & Johnson, Sanofi.

IB: None declared.

TCD: None declared.

JY: None declared.

CG: None declared.

JHM: None declared.

MG: Scientific advisory board/ speakers bureau: Almirall, Argenx, Biotest, Fresenius, GSK, Janssen, Leo Pharma, Lilly, Novartis, UCB.

PPS: Research support: Novartis, Abbvie, and Chugai. Travel grants or speaker fees: Abbvie, UCB, Janssen, Novartis

VO: Scientific advisory board/ speakers bureau: BMS, Johnson & Johnson, and UCB.

CH: None declared.

AS: research support/ clinical trials: Abbvie, Boehringer-Ingelheim, Celgene, Eli Lilly, Janssen-Cilag, LEO Pharma, Merck, Novartis, Pfizer; Scientific advisory board/ speakers bureau: Abbvie, Almirall, Hermal, Janssen, LEO, and Novartis, UCB.

Trial registration

German Clinical Trials Register (DRKS); Identifiers: DRKS00020755 (Registration Date: 2020-02-12), DRKS00020963 (Registration Date: 2020-04-09), DRKS00033790 (Registration Date: 2025-04-02), DRKS00037907 (Registration Date: 2025-09-15).

References

1. Ujiie, H. *et al.* Unmet Medical Needs in Chronic, Non-communicable Inflammatory Skin Diseases. *Front Med (Lausanne)* **9**, 875492 (2022). https://doi.org/10.3389/fmed.2022.875492
2. Blome, C. & Augustin, M. Measuring change in quality of life: bias in prospective and retrospective evaluation. *Value Health* **18**, 110-115 (2015). https://doi.org/10.1016/j.jval.2014.10.007
3. Kraft, R. *et al.* Global 10 year ecological momentary assessment and mobile sensing study on tinnitus and environmental sounds. *NPJ Digit Med* **8**, 162 (2025). https://doi.org/10.1038/s41746-025-01551-z
4. Seppen, B. F. Feasibility of Self-Monitoring Rheumatoid Arthritis With a Smartphone App: Results of Two Mixed-Methods Pilot Studies. *Journal of Medical Internet Research* **4** (2020-09-21). https://doi.org/10.2196/20165
5. Sahandi Far, M. *et al.* Cross-Platform Ecological Momentary Assessment App (JTrack-EMA+): Development and Usability Study. *J Med Internet Res* **27**, e51689 (2025). https://doi.org/10.2196/51689
6. Baker, A., Mitchell, E. J., Partlett, C. & Thomas, K. S. Evaluating the effect of weekly patient-reported symptom monitoring on trial outcomes: results of the Eczema Monitoring Online randomized controlled trial. *Br J Dermatol* **189**, 180-187 (2023). https://doi.org/10.1093/bjd/ljad163

7. Alinia, H. *et al.* Long-term adherence to topical psoriasis treatment can be abysmal: a 1-year randomized intervention study using objective electronic adherence monitoring. *British Journal of Dermatology* **176** (2017/03/01). https://doi.org/10.1111/bjd.15085
8. Yeboah, E. *et al.* Current trends in the application of causal inference methods to pooled longitudinal non-randomised data: a protocol for a methodological systematic review. *BMJ Open* **11**, e052969 (2021). https://doi.org/10.1136/bmjopen-2021-052969
9. Kumar, G. *et al.* Data Harmonization for Heterogeneous Datasets: A Systematic Literature Review. *Applied Sciences 2021, Vol. 11,* **11** (2021-09-06). https://doi.org/10.3390/app11178275
10. Ali, F. M. *et al.* A systematic review of the use of quality-of-life instruments in randomized controlled trials for psoriasis. *Br J Dermatol* **176**, 577-593 (2017). https://doi.org/10.1111/bjd.14788
11. Shikiar, R., Willian, M. K., Okun, M. M., Thompson, C. S. & Revicki, D. A. The validity and responsiveness of three quality of life measures in the assessment of psoriasis patients: results of a phase II study. *Health Qual Life Outcomes* **4**, 71 (2006). https://doi.org/10.1186/1477-7525-4-71
12. Bjelland, I., Dahl, A. A., Haug, T. T. & Neckelmann, D. The validity of the Hospital Anxiety and Depression Scale. An updated literature review. *J Psychosom Res* **52**, 69-77 (2002). https://doi.org/10.1016/s0022-3999(01)00296-3
13. Koopmann, A. *et al.* [Benefits of participatory involvement of patients in the development of a dermatological treatment app-A report from the practice]. *Dermatologie (Heidelb)* **75**, 562-565 (2024). https://doi.org/10.1007/s00105-024-05326-7
14. Domogalla, L. *et al.* Impact of an eHealth Smartphone App on the Mental Health of Patients With Psoriasis: Prospective Randomized Controlled Intervention Study. *JMIR Mhealth Uhealth* **9**, e28149 (2021). https://doi.org/10.2196/28149
15. Weigandt, W. A. *et al.* Impact of an eHealth Smartphone App on Quality of Life and Clinical Outcome of Patients With Hand and Foot Eczema: Prospective Randomized Controlled Intervention Study. *JMIR Mhealth Uhealth* **11**, e38506 (2023). https://doi.org/10.2196/38506
16. Gross, G. *et al.* Interdisciplinary approach to patients with psoriatic arthritis: a prospective, single-center cohort study. *Ther Adv Chronic Dis* **15**, 20406223241293698 (2024). https://doi.org/10.1177/20406223241293698
17. Mathews, S. C. *et al.* Digital health: a path to validation. *NPJ Digit Med* **2**, 38 (2019). https://doi.org/10.1038/s41746-019-0111-3
18. Amagai, S., Pila, S., Kaat, A. J., Nowinski, C. J. & Gershon, R. C. Challenges in Participant Engagement and Retention Using Mobile Health Apps: Literature Review. *J Med Internet Res* **24**, e35120 (2022). https://doi.org/10.2196/35120
19. Shiffman, S., Stone, A. A. & Hufford, M. R. Ecological momentary assessment. *Annu Rev Clin Psychol* **4**, 1-32 (2008). https://doi.org/10.1146/annurev.clinpsy.3.022806.091415
20. Kleiman, E. M. *et al.* Examination of real-time fluctuations in suicidal ideation and its risk factors: Results from two ecological momentary assessment studies. *J Abnorm Psychol* **126**, 726-738 (2017). https://doi.org/10.1037/abn0000273

21. Hall, R. *et al.* Development and Content Validation of Pruritus and Symptoms Assessment for Atopic Dermatitis (PSAAD) in Adolescents and Adults with Moderate-to-Severe AD. *Dermatol Ther (Heidelb)* **11**, 221-233 (2021). https://doi.org/10.1007/s13555-020-00474-9

Figures

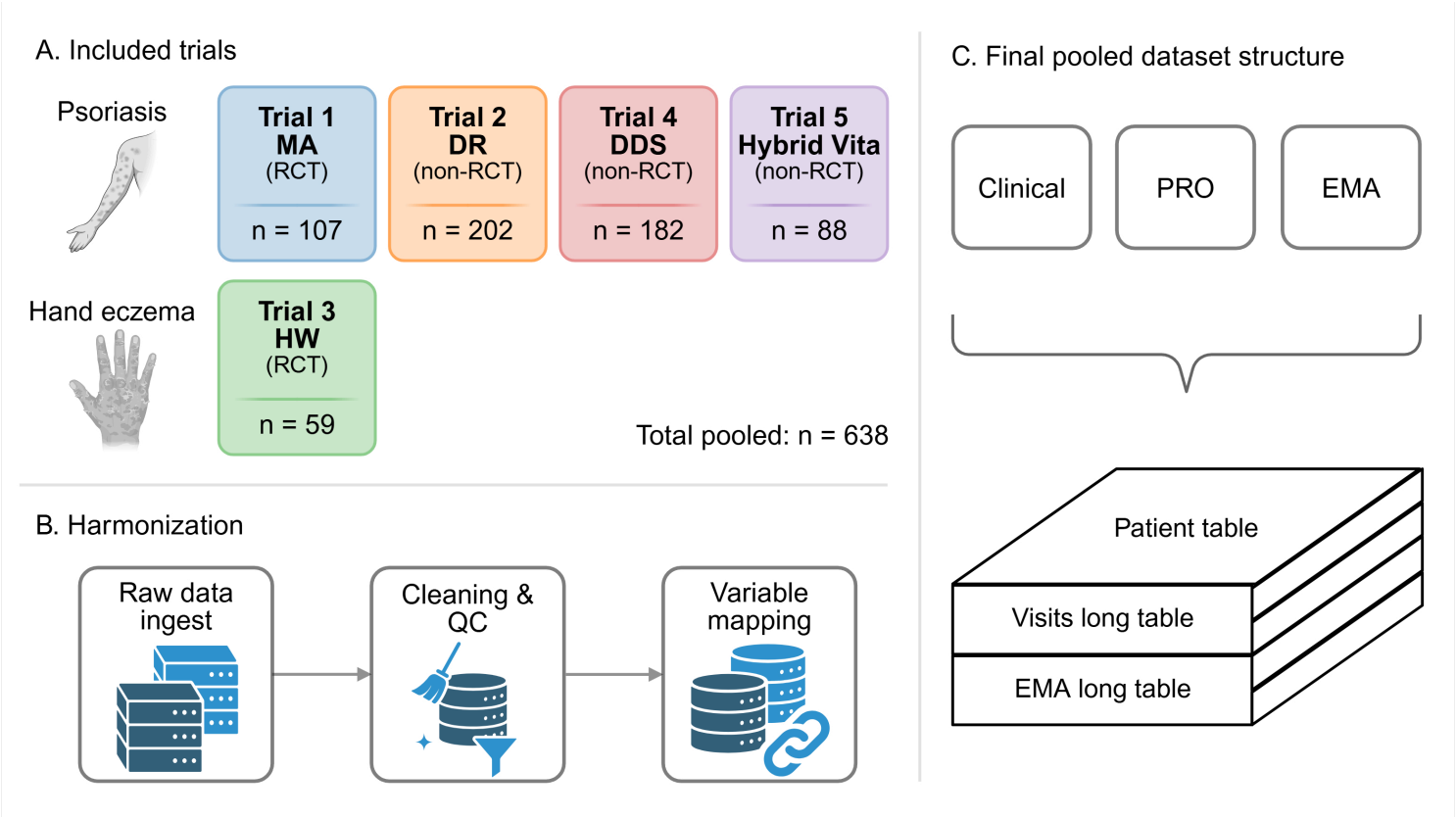


Figure 1

Pooled multi-trial dataset and harmonization workflow.

(A) Overview of the five contributing prospective dermatology clinical trials and their integration into a pooled analysis dataset. (B) Harmonization workflow including raw data ingestion, quality control, variable mapping, and alignment to a shared data dictionary. (C) Final harmonized data structure linking patient-level data with long-format clinical visit/PRO and EMA log tables enabling analyses of engagement, retention, and symptom trajectories.

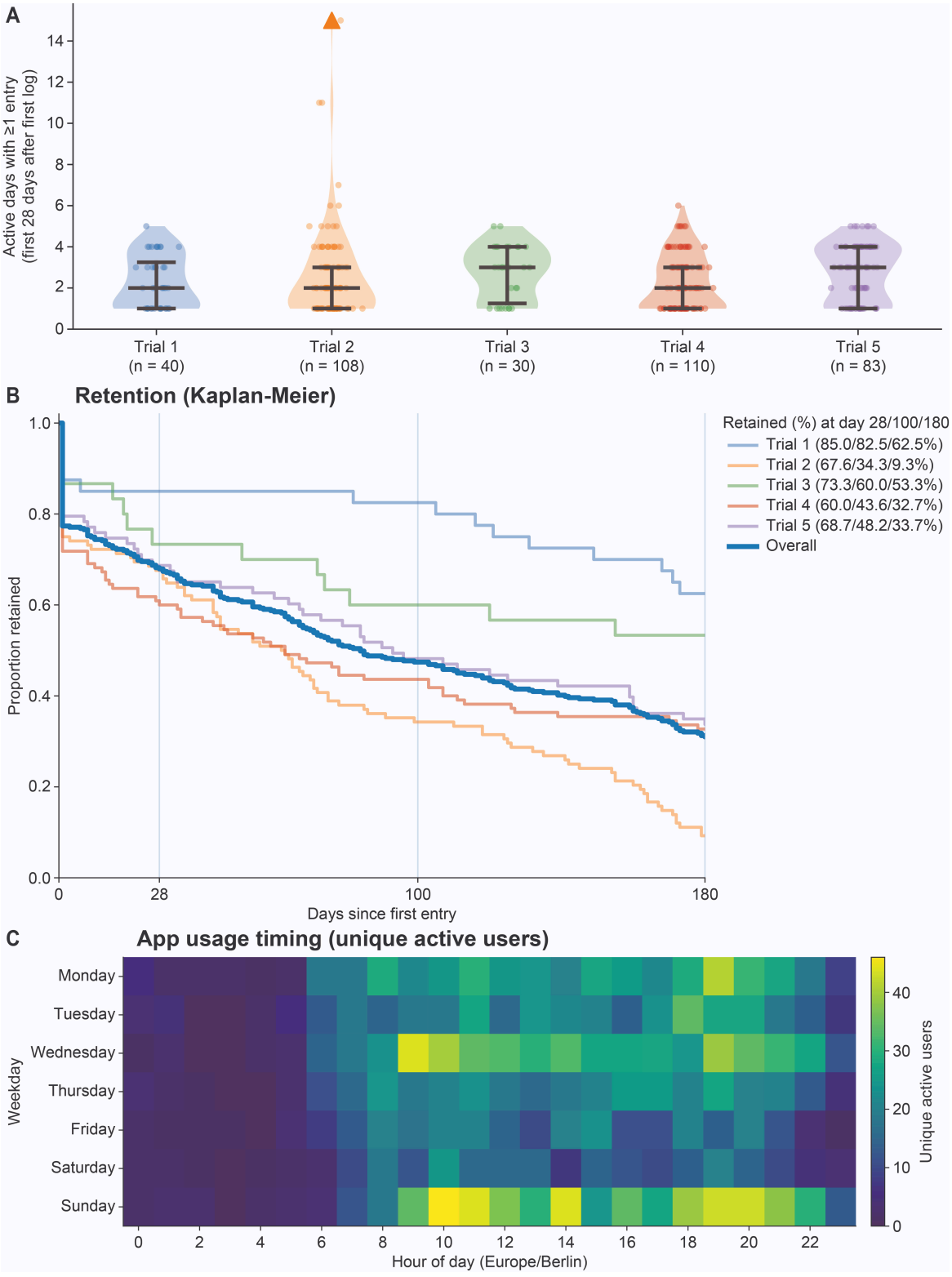


Figure 2

App engagement intensity, retention, and usage timing across trials.

(A) Distribution of early engagement defined as active days during the first 28 days after first activity, shown overall and by trial (trial-specific user counts are displayed in the panel). (B) Retention curves showing time from first to last observed activity day, overall and by trial. (C) Heatmap of aggregated

logging activity by day of week and hour of day (Europe/Berlin), illustrating diurnal and weekly usage patterns.

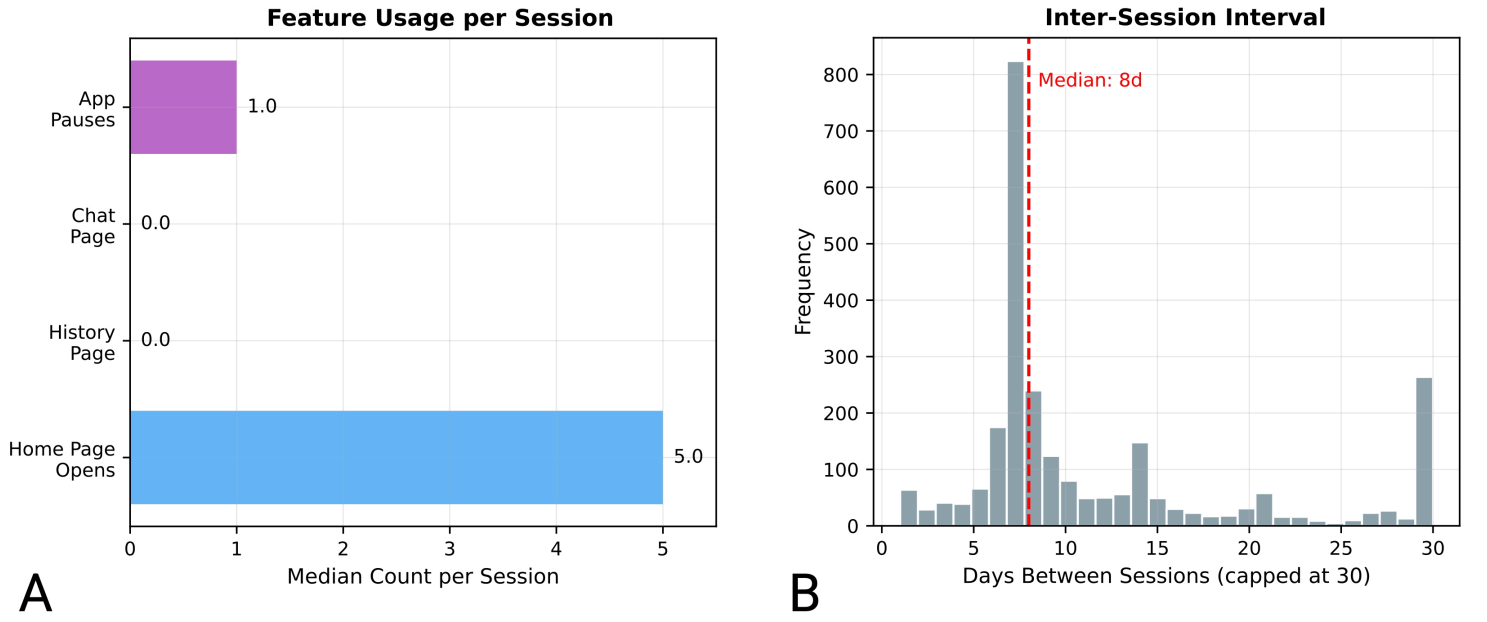


Figure 3

Session structure and revisit frequency.

(A) Median counts of selected in-app events per session. (B) Distribution of inter-session intervals (days between consecutive sessions within a user); dashed line indicates the median. For visualization, intervals are capped at 30 days.

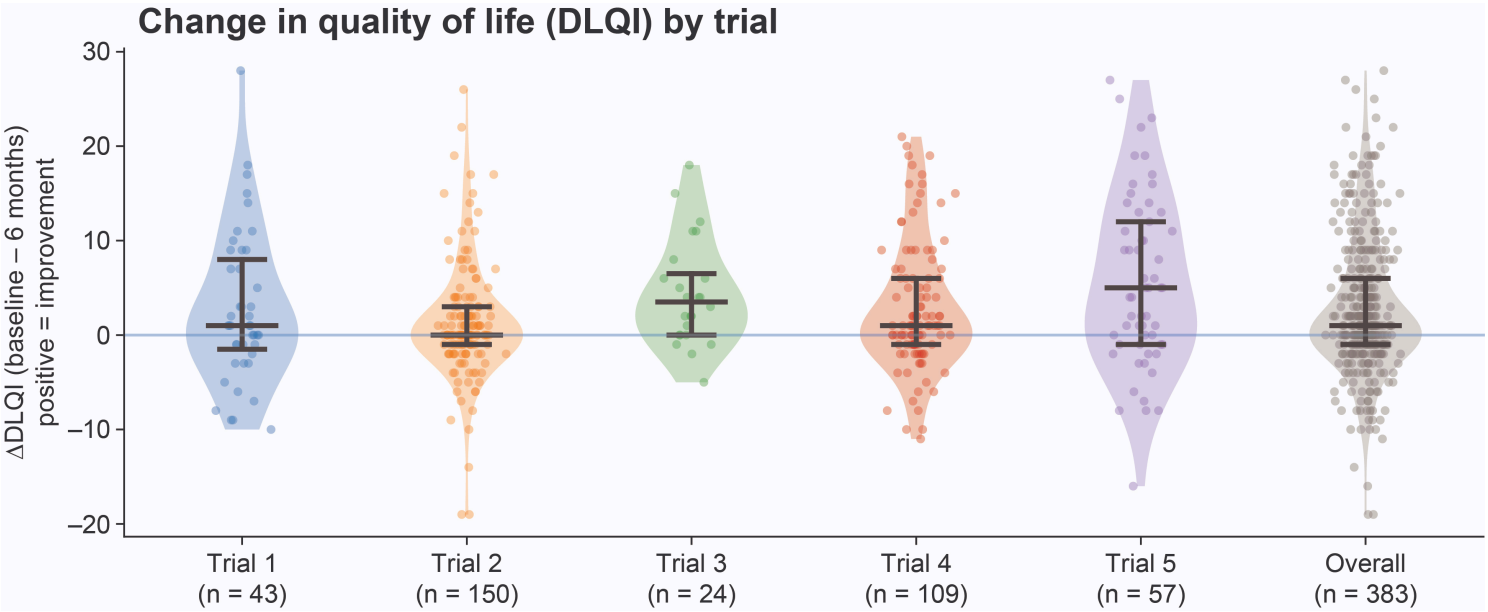


Figure 4

Clinical outcome change across trials.

Distribution of 6-month change in Dermatology Life Quality Index (ΔDLQI; baseline to 6 months), shown overall and by trial.

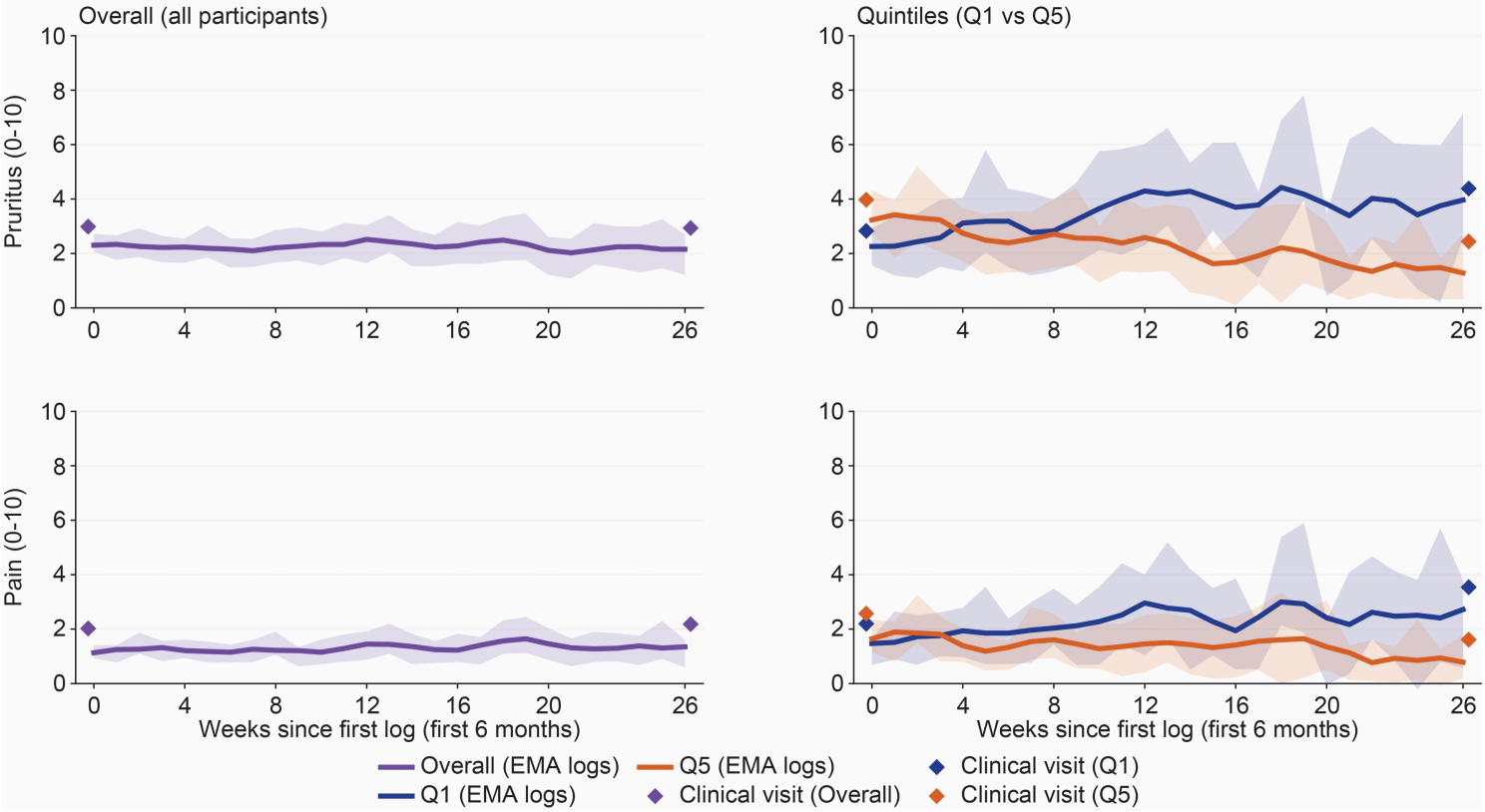


Figure 5

EMA symptom trajectories with clinical anchors and response stratification.

Weekly EMA trajectories for pruritus and pain over 26 weeks with clinic anchors at baseline and 6 months (left panels). Trajectories stratified by clinical response quintiles (ΔDLQI; Q1 = poorest response, Q5 = best response) show early and sustained separation between poorer and better responders (right panels).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementalFigureS1.pdf
- SupplementaryMaterial.docx