FITFUEL

AI-Enhanced Fitness App

Stakeholder Research Report

Discovery Phase · Draft for Review

Prepared by	Tim St Clair, Senior UX Lead
Date	June 2026
Version	Draft v1.0 — For Stakeholder Review
Status	Working Draft — Validation flags reference Appendix F
Source Basis	5 User Interviews · 2 Trainer Interviews · Market Analysis · Personas v4 · Journey Maps v3

RESEARCH DISCLAIMER: This report is based on primary qualitative research generated for the FitFuel UX course project. No personally identifiable information is included. Claims are rated by confidence level — all validation detail is in Appendix F and the FitFuel Source Validation Log v2. Internal use only.

01. Executive Summary

What this research found, what it means, and what comes next.

FitFuel is entering a fitness app market valued at $12–15 billion globally (2025), growing at 13–14% CAGR. The market is saturated at the surface but structurally underserved at the level of adaptive, life-aware coaching. Research across five user interviews, two practitioner interviews, a competitive market audit, and a validated source log confirms the claims below.

Three Confirmed Claims

- The gap is real. No current competitor has built a product that meets users after a lapse event, adapts to irregular real-world schedules, or serves the post-clinical return-to-exercise population at scale.
- The emotional design contract is non-negotiable. Five users and two trainers independently named shame mechanics — streaks, guilt notifications, missed-session commentary — as deletion triggers with clinical backing. Anti-shame design is the foundational product requirement.
- AI earns trust through scope honesty, not capability performance. Users and practitioners converge: AI is trusted when it explains its reasoning, acknowledges uncertainty, and stays within its competence.

Three Whitespace Opportunities

- Lapse recovery as a first-class product feature. No competitor has built this. Evidence: High.
- Context-aware scheduling — reasoning from calendar and life signals, not biometrics alone. Evidence: Medium-High.
- Mental wellness as coaching tone, not a feature category — movement reframed in mood and energy language. Evidence: Medium-High.

Research Constraints

- Sample (n=7) and selection bias limit generalisability. All findings are directional signals pending broader validation.
- Both trainers specialise in underserved, complex-constraint populations. Their observations do not represent the full fitness user spectrum.
- Design implications in this report are hypotheses, not validated solutions. All require testing before driving build decisions.

⚠ **Validation:** *See Appendix F — VN-00 — Full confidence ratings, source citations, and qualifiers for all claims in this report*

02. Project Context and Methodology

What FitFuel is building, why this research was conducted, and how.

2.1 Project Context

FitFuel is an AI-augmented fitness app designed to help users manage health goals, activities, and progress. The product accommodates fitness newcomers, returning routine-builders, and users with accessibility and disability needs. The core research question: Where does the current market fail users who want more than passive tracking but less than full human coaching?

2.2 Target Audiences

Audience	Description	Primary Design Challenge
A: Time-Constrained Professional	Fitness wrapped around dense schedules, travel, and variable energy. Former athletic identity common.	Plan-life mismatch: every app assumes a week they never have.
B: Returning Routine Builder	Has used fitness products before, dropped off when the plan stopped fitting their life, body, or confidence.	Re-entry shame and the absence of a 'return' product state.
C: User with Access Needs	Users with disability, injury recovery, vision impairment, chronic fatigue, or fitness culture anxiety.	Systematic exclusion: products are designed for a default user who is not them.

2.3 Research Sources

Source	Detail	Confidence Weight
User Interviews (n=5)	INT-01–INT-05, all three target audiences. ~60 min each. Semi-structured.	Primary — highest weight
Practitioner Interviews (n=2)	TRA-01: Strength coach, 12 yrs, Chicago. TRA-02: Adaptive coach, 18 yrs, Seattle.	Informed observation across hundreds of client relationships
Market Landscape Analysis	Competitive audit of 8–10 apps; market sizing from Grand View Research, Polaris, Business of Apps 2025.	Directional — market figures carry ±20% range
Source Validation Log v2	Structured confidence ratings across 6 design assumptions (High / Medium-High / Medium / Low-Medium / Low).	Authoritative reference for all claim status
Personas v4 + Journey Maps v3	Five user personas and four journey maps derived from interview synthesis.	Directional artifacts from interview data

2.4 Research Limitations

- n=7 is too small to support generalisable claims. Language implying universality should be read as directional.

- Selection bias: all participants are English-speaking, US-based, with prior fitness app experience. Satisfied users are not represented.
- Trainer testimony reflects specialist populations (complex-constraint, underserved), not mainstream gym-goers.

03. Market and Competitive Context

What the market looks like, where competitors are, and where the open space is.

3.1 Market Snapshot

$12–15B
Global Market 2025
App revenue · multiple analysts

13–14%
CAGR 2026–2033
Grand View Research · Polaris

3–8%
30-Day Retention Avg
Business of Apps 2025

25%
Revenue from Apps
Record share · up from prior years

⚠ **Validation:** *See Appendix F — VN-01 — Market size range, CAGR variance by analyst, and scope definition caveats*

3.2 Three Trends That Matter

Trend	What It Means for FitFuel	Confidence
AI adaptation is now baseline, not a differentiator	The question is no longer whether AI is present — it is how the AI reasons. Adapting to biometric signals alone is insufficient. Adapting to life context and lapse re-entry is the open design space.	Validated — June 2026 competitor audit
Market has shifted from acquisition to retention	Industry 30-day retention sits at 3–8%. Platforms gaining ground extend the relationship across lapse events, not feature volume. This directly favours FitFuel's positioning.	Partially validated — see VN-02
Mental wellness and fitness have converged — opportunity is in framing	Adding a wellness feature library would not differentiate FitFuel. The opportunity is at the language level: framing exercise in mood and energy terms across the entire coaching voice.	Partially validated — see VN-03

3.3 Two Underserved Segments

Segment A — The Irregular Exerciser

- Spans Time-Constrained Professionals and Returning Routine Builders. Core barrier is structural: no product accommodates a disrupted relationship with routine.
- UK Biobank data (~90,000 participants, Circulation 2024) confirms weekend-concentrated exercise produces comparable health outcomes to daily training — yet every major app treats it as a deviation to correct.
- Dropout mechanic confirmed: when a plan breaks, users abandon the app rather than adapt the plan.

Segment B — The User with Access Needs

- Users with disability, chronic illness, or fitness culture anxiety are systematically excluded by current product defaults.
- Activity Alliance research: 90% of users with impairments report being underserved by the app market.
- Design failure is structural: calorie deficits, step goals, and progressive load escalation are clinically contraindicated for some conditions.

⚠ **Validation:** *See Appendix F — VN-04 — Segment A weekend warrior citation; Segment B 90% underserved source; disability market CAGR variance*

3.4 Competitive Gap Analysis

Whitespace Opportunity	Competitor State (June 2026 Audit)	Evidence Strength
Lapse Recovery Architecture	No competitor (Fitbod, Freeletics, Future, Peloton IQ, Ray) detects an engagement gap, estimates detraining, and rebuilds a re-entry plan without framing the gap as failure.	High — BJH Psychology 2025 (n=58,881) + competitor audit
Context-Aware Scheduling	Current AI adapts to biometrics (HRV, sleep, load). None reason from schedule context: meeting overruns, travel days, compressed weeks.	Medium-High — UK Biobank validated; AI mechanism is directional
Mental Wellness as Coaching Tone	Wellness features are standard (Apple Fitness+, Peloton, Centr). None reframe exercise in mood/energy terms at the language level.	Medium-High — exercise–mental health link validated; tone mechanism is directional

04. User Insights

Eight themes from user interviews (INT-01–INT-05), tested against practitioner testimony (TRA-01, TRA-02). Themes marked BOTH appeared independently across both data sets — the strongest signals in this research.

Theme 1: Plan-Life Mismatch [BOTH — Highest Confidence]

The core friction is not motivation or knowledge — it is the structural mismatch between fixed app logic and unpredictable real lives. For users with chronic conditions (TRA-02), a fixed plan is not just inconvenient — it is medically incorrect.

> *Static plan in a dynamic life. That's the actual problem.*
> **INT-01, Time-Constrained Professional**

Design implication: A system that flexes around the user's actual week is the differentiating architecture. Calendar integration, energy signalling, and graceful degradation are the mechanisms.

Theme 2: Anti-Shame Design [BOTH — Highest Confidence]

Every participant raised, unprompted, the emotional harm caused by streak mechanics, guilt notifications, and missed-session commentary. TRA-01 frames the product consequence: shame causes users to hide behaviour from the app, breaking the AI feedback loop.

> *Be a friend who doesn't shame me. That's the energy.*
> **INT-01**

Design implication: Streaks off by default. Empty state shows tomorrow's session — never references the gap. All notification strings audited against a harm taxonomy before launch.

⚠ **Validation:** *See Appendix F — VN-05 — Uncertainty note: streaks work for some populations; implication is configurability, not blanket removal*

Theme 3: Decisive vs. Performative Personalisation [BOTH — High Confidence]

Participants distinguish sharply between cosmetic AI personalisation (name insertion, generic plan with a custom label) and AI that demonstrates actual knowledge of constraints and history. The former generates distrust and deletion.

> *Tell me the reasoning. 'Because you slept poorly, here's a lighter session.' That builds trust.*
> **INT-03, Returning Routine Builder**

Design implication: Every AI adaptation carries a plain-language rationale. Where AI is working from defaults, it says so. Onboarding promises personalisation that builds over time, not from day one.

⚠ **Validation:** *See Appendix F — VN-06 — Hypothesis: minimum personalisation threshold below which users prefer transparent generic plan over claimed-but-superficial custom*

Theme 4: Knowledge Is Not the Barrier [USER — High Confidence]

Most participants had prior athletic identities and strong fitness literacy. The barriers are structural: time, energy, life load, fear of re-injury, inability to moderate intensity. Apps that lead with education or simplified instruction alienate experienced users.

> *The information isn't the issue. Walk more, lift some, eat less garbage. The issue is I'm tired and I'm carrying a lot.*
> **INT-02**

Design implication: Onboarding asks 'What did your fitness life look like before?' not 'How fit are you now?' The lapsed athlete is a distinct archetype that beginner/intermediate/advanced framing misses entirely.

Theme 5: Permission to Do Less [BOTH — Highest Confidence]

All five users describe value in low-effort entry points on hard days. The critical framing: the low-bar option must be presented as 'this is what today calls for' — not as a consolation prize. TRA-01: every workout should have a 5-min, 15-min, and 45-min version.

> *The permission-to-do-less is the killer feature. The alternative: open the app, see a 45-min workout, close the app, feel bad.*
> **INT-03**

Design implication: Every plan state has a minimum viable version presented as the intelligent, appropriate choice — not partial credit.

Theme 6: Accountability Requires Realness [USER — High Confidence]

The most effective accountability every participant reported came from a real person. AI in a supporting role is valued if expectations are appropriately scoped. INT-03's Future experience shows the hybrid model is compelling but structurally fragile at subscription lapse.

> *The app is not my friend. I have friends. Lean into being useful, not into being a buddy.*
> **INT-02**

Design implication: AI scoped to logistics, scheduling, and data synthesis. Human escalation path as a feature, not a fallback. A human check-in tier is commercially testable — cost tolerance must be validated first.

Theme 7: Nutrition as Minefield [USER — High Confidence]

No participant wants calorie-centred defaults. Two have disordered eating histories for which logging is a direct psychological risk. INT-05 adds an accessibility dimension: food databases are unusable at speed with VoiceOver.

> *The absence of food tracking is a feature for me. Most apps treat it like a missing feature.*
> **INT-03**

Design implication: A prominent, genuine opt-out from nutrition tracking at onboarding. Educational mode offers simple principles without logging. Calorie counts not surfaced by default anywhere in the experience.

Theme 8: Inclusion by Default [USER — High Confidence]

INT-05 draws a precise distinction between an app that has an accessibility mode and one designed accessibly from the start. He identifies a structural AI problem: models trained on data that excludes disabled users produce wrong defaults before personalisation begins. Adjacent signals: INT-03 (neurodivergent design needs), INT-04 ('feel younger' framing as exclusion signal).

> *I want products where I'm the default for the parts of the experience that involve me.*
> **INT-05, User with Access Needs**

Design implication: Design the audio experience first. WCAG 2.2 AA is necessary but not sufficient. Hire disabled users as designers from day one. Accessibility features priced at zero.

4.2 Practitioner-Only Finding: AI Safety Boundary

Both trainers independently drew the same line: AI can do programming, scheduling, and pattern recognition. AI cannot do real-time interpretation, judgment under ambiguity, or clinical escalation. TRA-02 adds that AI's confident tone is itself a risk — users absorb AI mistakes without scepticism because confidence on a screen carries unearned authority.

> *AI doesn't know what it doesn't know. It speaks with confidence. The AI's mistakes are absorbed without question. That's a different kind of risk. It's more intimate.*
> **TRA-02, Adaptive Fitness Coach**

TRA-02's 8-point safe AI framework — the most actionable practitioner output in this research — is in Appendix G.

⚠ **Validation:** *See Appendix F — VN-07 — TRA-02 concern: AI in health is being deployed faster than it is being studied. No evidence base exists for what AI fitness coaching does to users over 6–24 months.*

05. Personas and Journey Maps

Synthesised user models derived from interview data. Directional artifacts — not validated profiles. Full persona cards and journey maps are in Personas v4 and Journey Maps v3 source documents.

5.1 Persona Summaries

Persona 1: The Calendar-Trapped Professional (INT-01)

32 · Sr. Product Marketing Mgr · Brooklyn, NY · ~$110K · Single

"Static plan in a dynamic life. That's the actual problem."

Goals	Pain Points	Key Needs
Get genuinely strong — one pull-up, fix back pain	Schedule disruption kills every plan — no graceful rebuild	Plan that survives travel without shame or catch-up logic
Manage stress through exercise	Guilt from apps leads to deletion, not motivation	No calorie tracking — hard limit due to eating history
A routine that becomes automatic	AI recommendations feel generic — 'nothing about me in it'	AI that actually uses calendar and sleep data
De-centre weight loss; strength is the real goal	Calorie features rarely offer a real opt-out	Social accountability layer; Pilates friend is the model

Persona 2: The Caregiver at the Edge (INT-02)

41 · Operations Manager (WFH) · Suburban Atlanta · ~$85K · Married, 2 kids + parent with Parkinson's

"I'm not lazy. I'm tired. There's a difference."

Goals	Pain Points	Key Needs
Lower pre-diabetic A1C and blood pressure	Time and energy never coexist — the window never opens	Home-only workouts — no gym commute
Strong enough to lift his mother safely	No Plan B — one-path apps fail him immediately	One button to press — zero cognitive friction
Find a fitness rhythm inside caregiving, not after it	Food logging incompatible with family cooking reality	Plan that adapts to caregiving emergencies without judgment
Fit old jeans eventually (fourth on the list)	One extra notification ping and the app is deleted	Progress in clinical language: A1C, blood pressure — not aesthetics

Persona 3: The All-or-Nothing Cycler (INT-03)

28 · Freelance Product Designer · East Austin, TX · ~$70K variable · ADHD, disordered eating history

"Choice is the enemy. Save me from myself, app."

Goals	Pain Points	Key Needs
Be a consistent mover — never achieved this	No middle gear — five sessions or zero	One session surfaced immediately — no library, no choices
Use exercise to manage ADHD cognition	Time blindness — misses sessions without noticing	AI that reads Oura + calendar and decides for him
Protect recovery from disordered eating history	Apps reinforce bad habits: streaks reward intensity sprints	Volume ceiling to protect against burnout cycles
Get pull-ups back from calisthenics era	Future worked at $150/mo but was unsustainable	Gender-neutral language throughout — non-negotiable

Persona 4: The Lapsed Athlete Finding Her Footing (INT-04)

52 · Elementary School Administrator · Minneapolis, MN · ~$75K · Married · Post-meniscus surgery, peri-menopausal

"The body I knew how to train is gone. Nobody handed me a manual for this one."

Goals	Pain Points	Key Needs
Feel confident in her body again — named before any functional goal	Fear of re-injury freezes her at every twinge	Clinical credibility — PT involvement before she trusts the AI
Walk-jog a 5K; play with future grandchildren	No app category fits — too experienced for beginner, not ready for athlete	Post-injury third category: not beginner, not athlete — disrupted
Reclaim athletic identity — will not surrender it	'Feel younger' framing signals the app wasn't built for her	Functional progress tracking: stairs, range of motion — not weight
Push back on perimenopause actively	Menopause invisible in mainstream fitness apps	Human escalation option for pain signals — AI cannot make that call

Persona 5: The Adaptive Mover (INT-05)

36 · Sr. Customer Support Specialist (remote) · Portland, OR · ~$52K · Legally blind, Long COVID

"I want products where I'm the default for the parts of the experience that involve me."

Goals	Pain Points	Key Needs
'Figure out how to be active without crashing' — his own problem statement	Every mainstream app is built for sighted users	Audio-first workout instructions throughout — no screen dependency

Rebuild muscle; carry groceries up four flights	Post-exertional malaise invisible to biometrics	AI that treats self-reported fatigue as primary data
Keep tandem cycling — main social and joy anchor	Accessibility settings reset on every app update	Accessibility preferences that persist across updates
Reach a 30–40 mile tandem ride; names the uncertainty directly	Streaks punish the medically correct decision to rest	Expressed AI uncertainty — confident wrong advice is a safety risk

5.2 Journey Map Highlights

Four journey maps were developed (Calendar-Trapped Professional, Caregiver at the Edge, Lapsed Athlete, All-or-Nothing Cycler). Three signals cut across all four maps:

- The plan must flex or the user disappears. Neither the professional nor the caregiver fails from low motivation. Both fail because a static plan meets a dynamic life. A rolling, adaptive plan is not a feature — it is the product.
- Re-entry is the most important — and most neglected — moment. The gap between a missed week and a deleted app is determined by what the product does when the user returns. An empty state showing tomorrow's session without referencing the gap is the minimum viable re-entry experience. No current competitor has built this systematically.
- Progress must be in the user's language. INT-01: back pain. INT-02: A1C. INT-03: pull-up. INT-04: stairs without thinking. None of them want streaks, body scores, or weight readouts. Progress must speak the language the user brought in.

06. Opportunity Areas

Three design opportunities grounded in research evidence. Each is stated with its design implication and a reference to full validation detail in Appendix F.

Opportunity 01: Lapse Recovery as a First-Class Product Feature

EVIDENCE HIGH

Every current fitness app is built for users who show up. None is built for users who stop and need to re-enter. A UCL and Loughborough University social listening study (n=58,881 posts, British Journal of Health Psychology 2025) documents that users abandon the app rather than adapt the plan when routines break — across all three target segments.

Design Implication:

An AI that detects lapse, estimates detraining by elapsed time, rebuilds the plan from current fitness state, and frames re-entry without surfacing the absence. Empty state shows tomorrow's session with no reference to the gap. Serves all three target segments.

⚠ **Validation:** *See Appendix F — VN-08 — Competitor gap confirmed by June 2026 feature audit. Behaviour documented by peer-reviewed social listening study. Re-entry AI mechanism is directional — requires prototype testing and clinical review before build.*

Opportunity 02: Context-Aware Scheduling

EVIDENCE MEDIUM-HIGH

Current AI adapts to the body. No current product adapts to the calendar. UK Biobank data (~90,000 participants, Circulation 2024) confirms that weekend-concentrated exercise produces physiologically equivalent health outcomes to daily training — yet every major app treats it as a deviation to correct.

Design Implication:

AI that reasons from scheduling signals in addition to biometric signals, redistributing weekly load across available windows without requiring the user to log every constraint manually. The weekend-concentrated exercise pattern is served as a first-class lifestyle, not corrected.

⚠ **Validation:** *See Appendix F — VN-09 — Weekend warrior equivalence validated (UK Biobank, Circulation 2024). AI calendar-reasoning mechanism is directional — requires interview validation and technical feasibility confirmation before specification.*

Opportunity 03: Mental Wellness as Coaching Tone

EVIDENCE MEDIUM-HIGH

Adding a meditation library does not differentiate FitFuel — Apple Fitness+, Peloton, Centr, Aaptiv, and Noom already include wellness content. A 2024 BMJ meta-analysis and a 2026 BJSM umbrella review confirm that walking, jogging, yoga, and strength training reduce depression symptoms across populations. The opportunity is framing exercise in these terms — not adding a breathing tab.

Design Implication:

AI microcopy and coaching tone that frames outcomes in mood and energy terms alongside physical metrics — across onboarding, plan summaries, re-entry moments, and achievement language. Lowest build cost of the three opportunities; broadest cross-segment applicability.

⚠ **Validation:** *See Appendix F — VN-10 — Exercise–mental health link validated (BMJ 2024; BJSM 2026). Tone-as-mechanism is directional — requires microcopy testing and user interview validation before voice guidelines are finalised.*

07. Recommended Next Steps

Nine actions required before design decisions are finalised. Ordered by dependency — items 1–3 are prerequisites for items 4–9.

#	Action	Purpose
01	**Broader user research sampling**	Recruit satisfied users, streak-motivated users, users for whom current apps work. Addresses the confirmation bias challenge from senior researcher review.
02	**Access-needs primary research**	Recruit 2+ participants beyond INT-05: chronic fatigue, mobility limitation, vision impairment. Single-source evidence for this segment must be expanded.
03	**AI trust threshold test**	Present mock AI recommendations to 8+ users across all three segments. Probe: act, modify, or ignore — and why? Tests the RQ3 disconfirming scenario.
04	**AI scope audit**	Map every planned AI output against the boundary both trainers drew: programming/scheduling (in scope) vs. interpretation, judgment, clinical escalation (out of scope).
05	**Language audit against trainer taxonomy**	Run every planned notification, empty state, missed-session message, and progress summary against the harmful language lists from TRA-01 and TRA-02.
06	**Post-clinical scope decision**	Determine whether the post-clinical population (discharged from PT, cardiac rehab, or oncology care) is within FitFuel's scope. If yes: define clinical advisory requirement. If no: document so the product does not inadvertently attract and fail them.
07	**Persona compression test**	Test whether INT-01 and INT-04 are better served by a single 'lapsed athlete' primary persona or two distinct ones. INT-02 diverges clearly on caregiving context.
08	**Nutrition opt-out framing test**	Test default-on tracking with opt-out vs. default-off with opt-in. Measure completion rate and 7-day retention. Hypothesis: default-off improves both.
09	**Source validation log update**	Promote all directional claims to Validated, Disconfirmed, or Retained-as-Hypothesis before broader distribution. FitFuel Source Validation Log v2 is the authoritative reference.

08. Appendix

A. Research Participant Overview

ID	Label	Age/Gender	Context	Audience	Key Constraints
INT-01	Calendar-Trapped Professional	32 / F	Sr. Product Mktg Mgr, Brooklyn NY	Time-Constrained Professional	Unpredictable schedule, SSRI, lower back pain, disordered eating history
INT-02	Caregiver at the Edge	41 / M	Operations Manager, Atlanta	Time-Constrained (Caregiver Variant)	Parent w/ Parkinson's, 2 kids, pre-diabetic A1C, football knee, 6 hrs sleep
INT-03	All-or-Nothing Cycler	28 / M	Freelance Designer, Austin TX	Returning Routine Builder	ADHD, disordered eating history, variable income, gym-culture aversion
INT-04	Lapsed Athlete Finding Footing	52 / F	Elementary Admin, Minneapolis	Returning Routine Builder	Post-meniscus surgery, peri-menopausal, reading glasses, 20-yr fitness identity
INT-05	Adaptive Mover	36 / M	Customer Support, Portland OR	User with Access Needs	Legally blind (RP), Long COVID / post-exertional malaise, budget-limited
TRA-01	Strength Coach	38 / F	NSCA-CPT, private studio, Chicago	Practitioner	12 yrs exp; busy professionals 35–55; ACL recovery background
TRA-02	Adaptive Coach	46 / NB	ACE-CPT, adaptive coaching, Seattle	Practitioner	18 yrs exp; chronic illness, disability, post-surgical; sliding-scale pricing

B. Design Assumption Confidence Summary

Assumption	Confidence	Key Risk / Qualifier
A-01: Structural attrition creates a re-entry market	HIGH	Evidence confirms attrition. Does not confirm users actively seek a re-entry product — must come from primary interviews.
A-02: Users want adaptive AI that adjusts over time	MEDIUM–HIGH	JMIR RCT tested reinforcement learning, not conversational AI coaching. Trust handoff is the unresolved design question.
A-03: Apple is not a near-term platform threat	MEDIUM	Based on journalism (Bloomberg, Feb 2026), not official Apple statement. Incremental health releases may collectively substitute.
A-04: Disability fitness segment worth designing for from the start	MEDIUM	Market size figures vary up to 5x across analysts. Accessibility is a founding design constraint regardless of market sizing.

A-05: Biometric-grounded AI with strong directional POV is the right model	MEDIUM–HIGH	No FitFuel-specific user testing conducted. Trust handoff — follow AI vs. override — is the key unresolved design question.
A-06: GLP-1 users represent an emerging sub-persona	LOW–MEDIUM	Noom-internal observational data only. Horizon signal, not a validated design requirement.

C. Senior Researcher Pushbacks

Pushback	What It Would Change
Sample size doesn't support generalisability (n=7, all US-based, prior fitness experience only)	All findings should be framed as directional signals. Language implying universality must be qualified throughout.
Confirmation bias — sample selected for users with frustration, not satisfaction	Analysis needs engagement with the counterfactual: what does the person this product is NOT for look like?
Trainer testimony treated as expert validation rather than a second perspective	Trainer observations should be explicitly scoped to their specialist populations. The absence of disagreeing trainers is not evidence they're wrong.
Design implications leap from insight to solution without showing inferential steps	Design implications should be framed as hypotheses to test, with explicit failure modes named.
Emotional tone framing is editorial, not analytical	Tone characterisations should be labelled as analyst interpretation, not findings.

D. Research Question Response Status

Research Question	Assessment
RQ1: Where do existing apps fail users who want more than tracking but less than full coaching?	CONFIRMED. Gap is real and larger than framed. Post-clinical void extends scope beyond the brief.
RQ2: What triggers cause disengagement, and which are AI-addressable?	PARTIALLY CONFIRMED. Schedule disruption, energy misalignment, boom-bust intensity are addressable. Cost and physical limitation are not.
RQ3: Under what conditions do users act on AI recommendations?	DIRECTIONAL. Trust increases with explained reasoning and honoured prior constraints. Alarming inverse: users may comply with confident-sounding wrong AI advice.
RQ4: Do three audiences share core barriers or have distinct profiles?	CONFIRMED — SHARED AND DISTINCT. Shared: anti-shame, permission to do less, no calorie defaults, real personalisation. Distinct at constraint type and clinical profile.
RQ5: What makes the Access Needs user feel safe — or excluded?	CONFIRMED DISTINCT. Barriers are structurally different. WCAG 2.2 AA is necessary but not sufficient.

E. Source Key

Claim	Source Reference
Weekend warrior health equivalence	UK Biobank, Circulation 2024, ~90,000 participants
30-day retention benchmarks (3–8%)	Business of Apps 2025
Market size ($12–15B, 13–14% CAGR)	Grand View Research + Polaris Market Research (directional range)
Lapse / social listening study	UCL + Loughborough, n=58,881 posts, British Journal of Health Psychology 2025
Exercise + mental health meta-analysis	BMJ 2024 + British Journal of Sports Medicine 2026 umbrella review
mHealth fitness app attrition	JMIR mHealth Jan 2026, n=719,421
AI adaptive prescription RCT	Doherty et al., 2024, JMIR mHealth DOI: 10.2196/49443
Disability / access needs underserved	Activity Alliance / Accessercise 2023
GLP-1 user behaviour in fitness apps	Noom corporate blog, Jan 2026, n=30,239 (observational, not independently replicated)
Apple Project Mulberry shelved	Bloomberg / Mark Gurman, Feb 2026 (journalism — not an official Apple statement)

F. Validation Notes (Referenced in Body of Report)

All ⚠ Validation annotations in the report body reference the entries below. These contain the full confidence ratings, qualifiers, and action requirements that were summarised inline.

VN-00 — General Confidence Framework

⚠ HIGH: Supported by peer-reviewed primary research or verified government source — safe to cite with attribution. MEDIUM–HIGH: Peer-reviewed with methodological caveats — cite with qualification. MEDIUM: Corroborated secondary sources — directionally credible, requires primary validation before design decisions are finalised. LOW–MEDIUM: Single secondary or proprietary source — do not cite as evidence in stakeholder documents. LOW: Unverified or actively contested — discard for citation.

VN-01 — Market Size and CAGR

⚠ Market size figures are consistent across Grand View Research, Polaris, and Business of Apps when scoped to app-only revenue. CAGR projections vary 11–15% across analysts by scope definition. Do not present a single number without noting the range.

VN-02 — Retention and Re-Engagement Positioning

⚠ 30-day retention benchmarks (3–8%) confirmed by Business of Apps 2025. Competitive advantage of re-engagement architecture for FitFuel is directional — supported by industry analysis but not independently quantified for a new entrant.

VN-03 — Mental Wellness Convergence

⚠ Market convergence documented (Feed.fm 2026 Digital Fitness Ecosystem Report). Exercise–mental health link validated by peer-reviewed meta-analyses (BMJ 2024; BJSM 2026). Tone-as-mechanism recommendation is directional — requires microcopy testing and user interview validation before it becomes a voice guideline.

VN-04 — Underserved Segments

⚠ Segment A: Weekend warrior health equivalence validated (UK Biobank, Circulation 2024). Segment B: 90% underserved figure from Activity Alliance / Accessercise 2023. Disability fitness market CAGR figures vary up to 5x across analysts (5.96%–22.4%) depending on whether software-only or adaptive physical equipment is included. Do not use specific market size figures without noting the range and methodology uncertainty.

VN-05 — Anti-Shame Design

⚠ Anti-shame findings are convergent across all 7 participants. Uncertainty: this sample selected for users with prior frustration. Streaks and intensity mechanics work for some populations — users motivated by completion metrics, competitive users, and high-frequency gym-goers are not represented. The implication is configurability and smart defaults, not blanket removal of streak mechanics from the product.

VN-06 — Personalisation Trust Threshold

⚠ Hypothesis only — not validated by this research. There may be a minimum personalisation threshold below which users prefer a transparent generic plan over a claimed-but-superficial custom one. This requires direct testing: present users with (a) a labelled-as-custom but generic plan and (b) a transparently generic plan. Measure trust and engagement.

VN-07 — AI Safety in Health

⚠ TRA-02 raises a concern TRA-01 does not: AI in health is being deployed faster than it is being studied. There is no peer-reviewed evidence base for what AI fitness coaching does to users over 6 months or 2 years. This is not a reason not to build — it is a reason for epistemic humility in how the product presents itself, and for maintaining conservative defaults throughout.

VN-08 — Lapse Recovery Architecture

⚠ Competitor gap confirmed by June 2026 feature audit across eight products. Lapse-abandonment behaviour documented by peer-reviewed social listening study (UCL + Loughborough, n=58,881, British Journal of Health Psychology 2025). The specific product mechanism — how the AI estimates detraining by elapsed time and rebuilds the plan — is a directional hypothesis. It must be validated through prototype testing and clinical review before it is specified for build.

VN-09 — Context-Aware Scheduling

⚠ Weekend warrior health equivalence validated (UK Biobank, Circulation 2024). Schedule-driven dropout documented across interviews. The AI calendar-reasoning mechanism is directional — reasoning from calendar signals rather than biometric signals alone has no direct precedent in the current competitor set and requires both interview validation (do users trust schedule-adaptive recommendations?) and technical feasibility confirmation before it is specified.

VN-10 — Mental Wellness as Coaching Tone

⚠️ Exercise–mental health link validated by peer-reviewed meta-analyses (BMJ 2024; BJSM 2026 umbrella review). Tone-as-mechanism is directional. The claim that framing exercise in mood and energy terms produces better retention or engagement than standard framing has not been tested. This requires A/B microcopy testing and user interview validation before voice guidelines are finalised. The build cost is low relative to the other two opportunities, making this a low-risk place to test the framing hypothesis early.

G. TRA-02 Safe AI Framework

The following eight requirements were generated by TRA-02 (Adaptive Fitness Coach) from 18 years of experience coaching disabled, chronically ill, and post-clinical populations. This is the most actionable practitioner output in this research. Every planned AI feature should be evaluated against this framework before build.

- Explicit refusal when beyond capability — the AI must know when to stop and say so.
- Conservative defaults throughout — when in doubt, prescribe less. The damage from too much exercise is greater than the damage from too little for most populations in this research.
- Radical transparency about training data and population coverage — users should be able to see what the AI was trained on and who it was calibrated for.
- Clear separation: AI does programming and scheduling. Humans do interpretation, judgment, and clinical escalation. These roles should never blur.
- Subjective input treated as primary data, not secondary to biometrics. Self-reported fatigue, pain, and readiness are more reliable than HRV for populations with chronic conditions.
- Escalation to human practitioners as a feature, not an exception. The pathway from AI to human should be visible, frictionless, and not framed as a failure of the AI.
- Expressed uncertainty in the UI — 'I'm not sure' is more trustworthy than confident-sounding wrong output. The product should model appropriate epistemic humility.
- User-visible and correctable model of what the AI has learned — users should be able to see, correct, and reset what the AI believes about them.