

◆ THE QUESTION WE SET OUT TO ANSWER ◆

Can routine outcome monitoring — questionnaires only, no session text — be turned into a personalized forecast of who improves, in which domain, and who is at risk of deterioration?

We set out to answer this with **11 years** of the clinic's routine data — a broad, multi-informant, session-by-session battery.

OUR APPROACH a two-phase machine-learning program — **benchmark** a per-client outcome forecast at **session 3-5**, then **personalize** it by domain & flag deterioration

17,326

sessions

924

dyads

693

clients

504

therapists

129,894

questionnaire administrations

22

measures

3

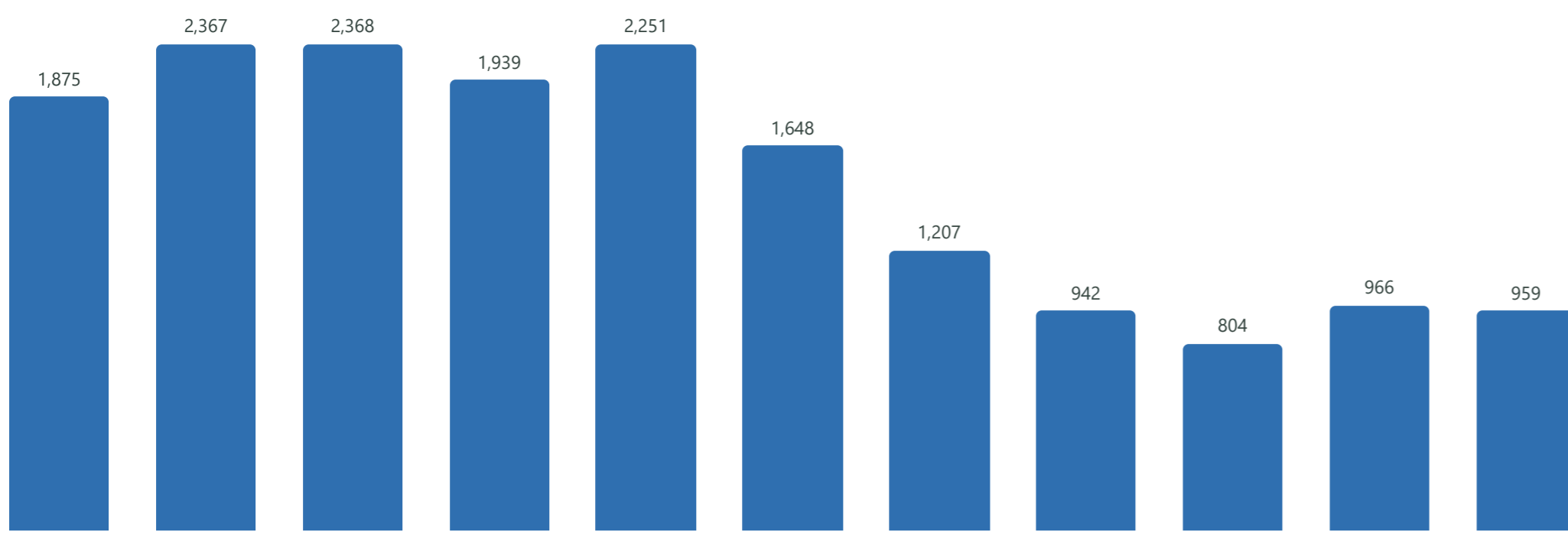
informants

11

years 2013-24

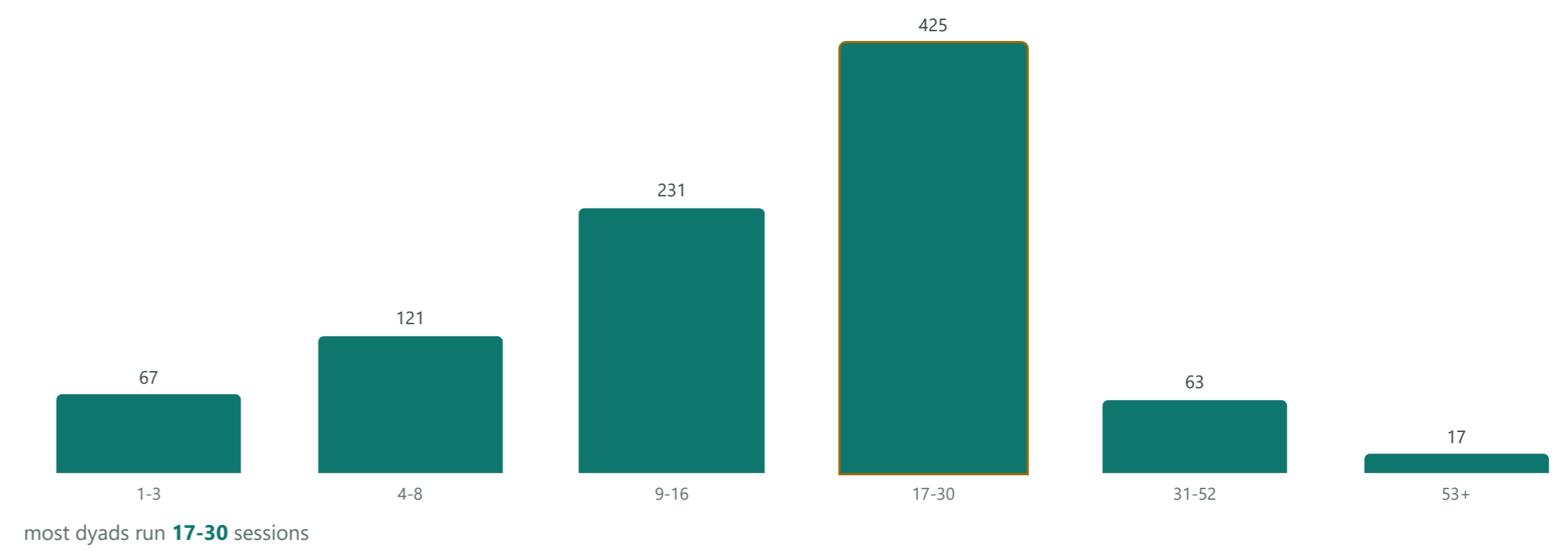
Sessions per year

17,326 across 11 years (COVID/war dip visible)



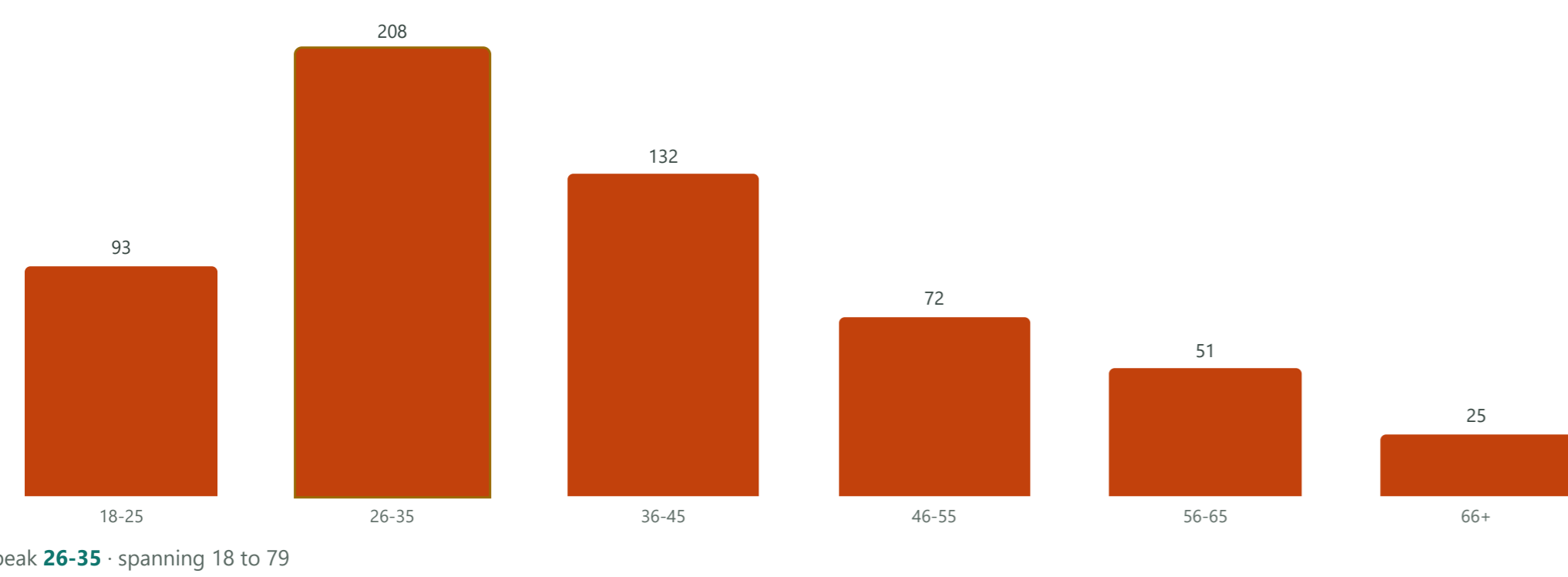
Treatment length

sessions per dyad · median 18 (IQR 10-24)



Client age distribution

consented clients · broad adult range



Three informants

same constructs, different eyes (dyads with data)



enables agreement & empathic-accuracy features

Who the clients are

consented analytic sample

37.9

mean age (SD 13.2)

59.6%

female

18-79

age range

a broad adult outpatient population

Returning clients

clients seen across multiple academic years

10

max years, one client

412

dyads with pre & post OQ-45

1.5

mean clients per therapist/yr

The measurement battery — 22 instruments, full names

informant (dot): client / therapist / therapist-on-client · items · consented N people

SESSION

before / after every session

● ORS	Outcome Rating Scale	4 it	699
● HSCL	Hopkins Symptom Checklist	11 it	699
● WAI-6	Working Alliance Inventory	6 it	622
● POMS	Profile of Mood States	12 it	621
● SES	Session Evaluation Scale	1 it	409
● Rupture	Rupture self-report item	1 it	622

The longitudinal core — repeated before/after every session, giving each client a dense within-treatment trajectory.

BASELINE

once per academic year

● OQ-45	Outcome Questionnaire-45	45 it	599
● BDI	Beck Depression Inventory	21 it	598
● IIP	Inventory of Interpersonal Problems	32 it	593
● DERS-18	Difficulties in Emotion Regulation Scale	18 it	560
● ERQ	Emotion Regulation Questionnaire	10 it	497
● PID-5	Personality Inventory for DSM-5	25 it	347

BASELINE (CONT.)

once per academic year

● SPIN	Social Phobia Inventory	17 it	448
● SE	Rosenberg Self-Esteem Scale	10 it	449
● SWLS	Satisfaction With Life Scale	5 it	326
● SHEEHAN	Sheehan Disability Scale	3 it	189
● IPDE	Int'l Personality Disorder Examination	77 it	143

A once-a-year intake profile — personality, symptoms, emotion regulation and well-being — the rich starting point each forecast builds on.

THERAPIST

self-report

● WAI-6 (T)	Working Alliance Inventory (therapist)	6 it	600
● IIP (T)	Inventory of Interpersonal Problems (therapist)	32 it	374
● DERS (T)	Difficulties in Emotion Regulation (therapist)	18 it	328

The therapist's own interpersonal style & emotion regulation — the basis for therapist-effect predictors.

CROSS-RATER

therapist-on-client

● tc-ORS	Outcome Rating Scale (therapist-rated)	4 it	522
● tc-POMS	Profile of Mood States (therapist-rated)	12 it	482

The therapist's view of the same client — enables client-therapist agreement & empathic-accuracy features, among the strongest known predictors.

HOW WE WILL ANSWER IT

a two-phase machine-learning program · questionnaires only · battery feeds the model

