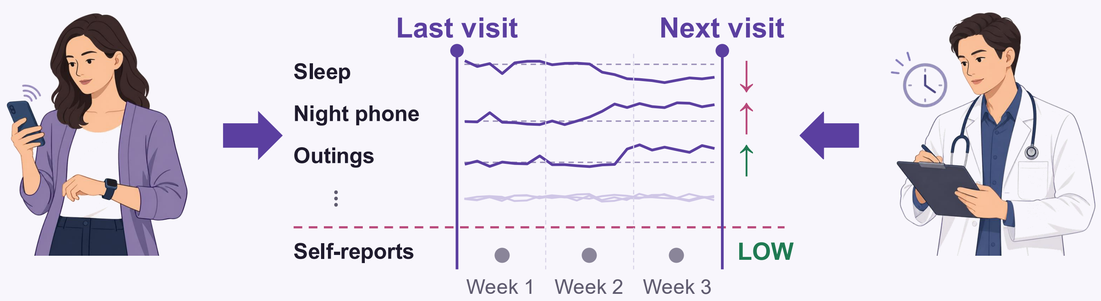
Designing Clinician-Centered Explanations for Sensing-Based Longitudinal Mental Health Monitoring

Esther Kang, Youngji Koh, Seungwan Jin, Uichin Lee
esther.kang@kaist.ac.kr

01 | MOTIVATION

Weeks of patient data, minutes to make sense of it



Passive sensing and EMA accumulate between visits, yet the **signals often conflict** and a consultation lasts only minutes. Psychiatrists must quickly decide **what matters, what supports it, and what remains uncertain**.

Three XAI gaps in supporting clinical reasoning

Each signal is only partial
Self-reports and sensed behavior each show only part of the patient

Too little vs. too much
Too little evidence to judge an interpretation; too much to read in a short visit

Presenting ≠ reasoning
Prior dashboards focus on presenting insights, less on how clinicians weigh them

Research question *How should AI interpretations of sensing data be explained so that clinicians can reason with them, rather than simply accept or reject them?*

02 | DESIGN REQUIREMENTS

We interviewed 12 psychiatrists in South Korea about how they read patient data and what explanations they need. Three requirements emerged:

DR1 Rapid Understanding
With limited time, clinicians preferred to see the main concerns first and access details only when needed.

DR2 Evidence-Grounded Interpretation
Clinicians wanted AI interpretations to remain traceable to the evidence behind them, including conflicting evidence.

DR3 Support for Clinical Inquiry
Clinicians valued follow-up questions that helped clarify what remained uncertain before making a judgment.

03 | MULTI-AGENT REASONING

Patient-generated health data
Sleep · Activity · Phone use · Self-report · + more

↓

Preprocessing
features vs. the individual baseline, across time windows

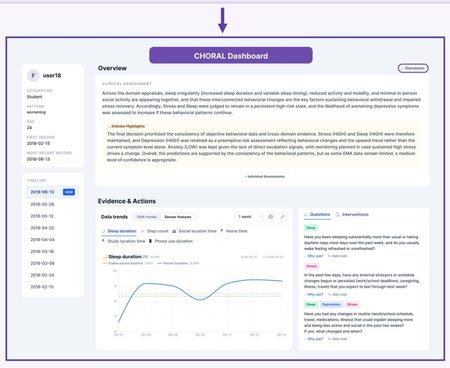
↓

Depression agent · **Anxiety** agent · **Stress** agent · **Sleep** agent

↓

Moderator agent
synthesizes the agents' assessments

↓

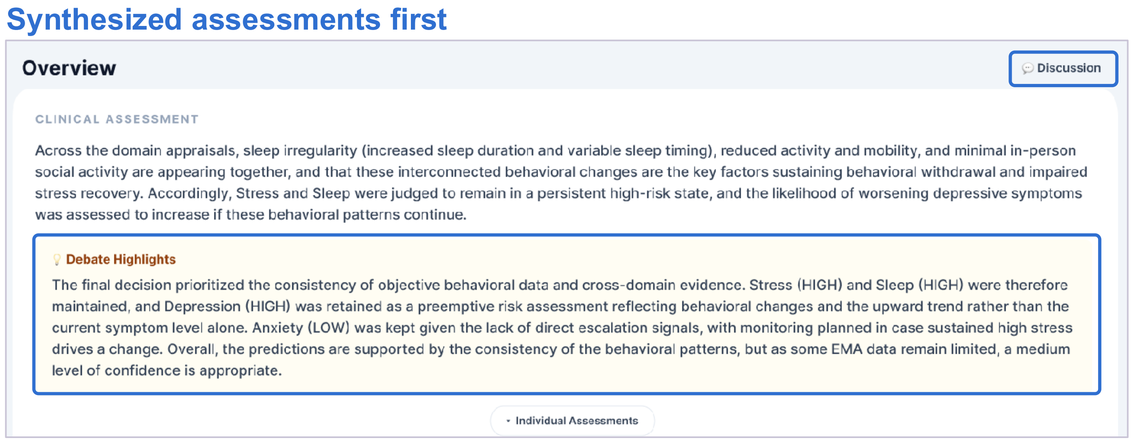


* Prototype running on the CES dataset

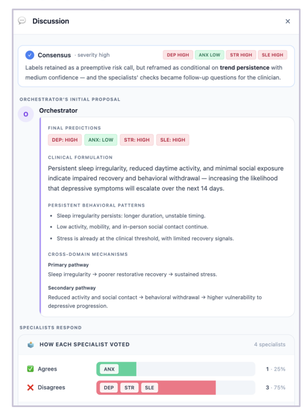
04 | THE EXPLANATION INTERFACE

A Overview: synthesis first, rationale on demand — DR1

Synthesized assessments first



Details on demand



Synthesis first (DR1)

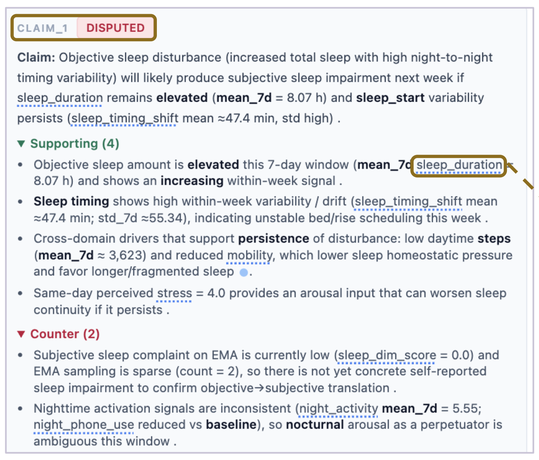
- The synthesized assessment highlights key changes
- Debate highlights show how it was formed
- The overview fits one screen for a quick first read

Rationale on demand (DR1)

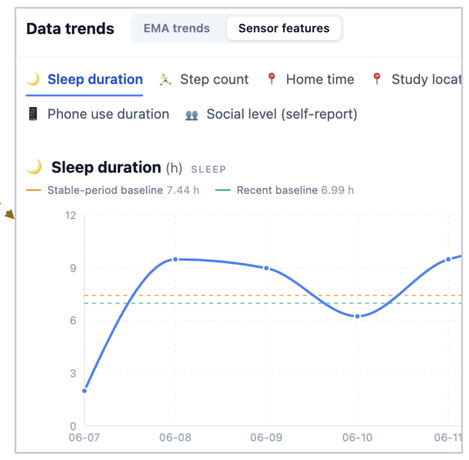
- Agent assessments expand below the overview
- Full discussion: consensus and each specialist's vote
- Detailed reasoning is available on demand, not by default

B Evidence & Actions: from data trends to clinical inquiry — DR2 DR3

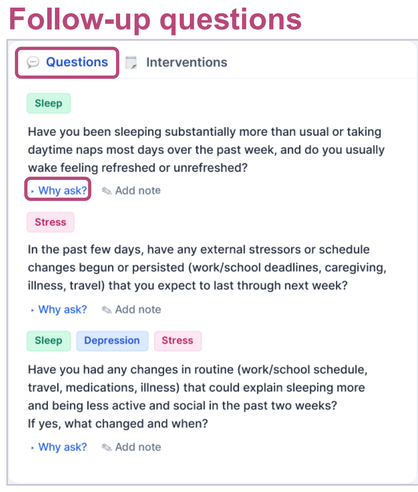
Claim with evidence



Linked data trends



Follow-up questions



Evidence-grounded (DR2)

- Competing interpretations remain visible
- Each claim shows supporting and counter evidence
- Evidence links back to the longitudinal data

Clinical inquiry (DR3)

- Questions are generated from unresolved findings
- 'Why ask?' shows the uncertainty each one resolves
- Interventions connect findings to next steps

05 | KEY IMPLICATIONS & FUTURE DIRECTIONS

IDENTIFY

Lead with what matters
Prioritized findings come first; the detailed rationale stays one step away so a brief visit is not spent reading.

EVALUATE

Show the evidence, keep the disagreement
Supporting and counter evidence tied to individual baselines let clinicians judge an interpretation rather than accept it.

ACT

Turn uncertainty into questions
What the data cannot resolve becomes patient-specific questions and grounded intervention ideas for the consultation.

FUTURE DIRECTIONS

- **Explain the data itself:** sensing quality, missing data, life context
- **Evaluate in real consultations:** practicing psychiatrists and clinical patients
- **Make explanation interactive:** clinicians correct it, the system learns from it