AI Chatbot-Enabled Patient Group Management in Rare Diseases: Outcomes from a Real-World WeChat-Based Deployment in China

Adele Li, MBA, David Wang, MBA, Yixuan Zhou, MSc.

MSR16

Background

Rare disease patients in China face compounded burdens: fragmented patient advocacy group (PAG) infrastructure, limited 24/7 support access, low health literacy and significant geographic barriers to specialist care.

The pressure is structural rather than attitudinal. Hospital outpatient services operate at high load, leaving little room for the one-to-one explanation that patients need after a diagnosis they have never heard of. Peer communities have filled the gap, but they depend on volunteers whose availability is bounded and whose answers vary.

The result is information asymmetry at precisely the moments when patients are least able to tolerate it — at diagnosis, at recurrence, and at the point where a side effect prompts a private decision to stop treatment.

Objective

To evaluate the clinical and operational outcomes of an AI chatbot deployed within WeChat-based patient communities, assessing its impact on patient support scalability, information quality, medication adherence, and the generation of real-world patient insights for healthcare decision-making.

Methods

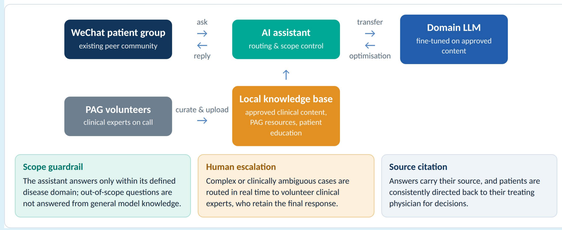
An HEOR impact case methodology was applied to a documented real-world deployment of an AI chatbot ("Panda and Friends") within patient WeChat groups, operated in partnership with the China Organisation for Rare Disorders (CORD).

The architecture uses a disease-specific LLM fine-tuned on approved clinical content, PAG resources and validated patient education materials, with a human volunteer expert escalation protocol for complex cases and a scope guardrail restricting the assistant to its defined disease domain.

Performance metrics captured daily question-and-answer volume, response accuracy, user reach, escalation rates, and downstream patient actions including specialist referrals and adherence improvement. Operating data covered 232 consecutive days (1 January – 20 August 2025). Contextual benchmarking used 2025 DLP physician survey data (n=4,500) on patient AI usage patterns.

Figure 1. Operating architecture — AI assistant embedded inside the patient community

The chatbot sits within existing WeChat peer groups rather than requiring patients to adopt a new platform



Contextual benchmarking against the 2025 DLP physician survey (n=4,500) situates the deployment against how physicians report their patients already using AI — behaviour occurring with or without a governed channel.

Results

The deployed system processed approximately 280 patient Q&A sessions per day at zero incremental staffing cost, providing 24/7 continuous support unavailable through conventional volunteer-staffed PAG models. Full operating data across 232 consecutive days record 72,549 sessions, a mean of 313 per day (median 306, range 108–545).

Figure 2. Sustained daily question-and-answer volume over 232 consecutive days

1 January – 20 August 2025; every day of operation, no sampling

72,549	313	232	24/7
total Q&A sessions handled	mean sessions per day (median 306)	consecutive days without service gap	availability at zero incremental staffing cost

Volume is continuous rather than clustered around clinic hours. The dips are seasonal — national holidays and the summer period — not service interruptions, which is the operational claim a volunteer-staffed model cannot make.

Volume held without a single service gap across the full period. The seasonal dips track national holidays, which is when volunteer-staffed support is least available and patient anxiety is least likely to wait — the coverage asymmetry is the operational point.

Figure 3. Conventional volunteer-staffed PAG support versus AI-enabled support

What changes is not the presence of support, but its availability, consistency and marginal cost

	Conventional model: Volunteer-staffed peer group	AI-enabled model: Chatbot with expert escalation
Availability	Bounded by volunteer availability; overnight and holiday questions wait	Continuous 24/7 response, including at diagnosis and recurrence — when distress peaks
Capacity	Scales linearly with volunteer headcount — the binding constraint	~313 sessions/day absorbed at zero incremental staffing cost
Consistency	Answer quality varies with which volunteer is present	Single curated knowledge base; every answer sourced
Marginal cost	Each additional group requires additional trained people	Additional groups add compute, not headcount
Data capture	Conversations are unstructured and effectively unanalysable	Structured interaction log becomes an analysable evidence asset

Note Volunteers are not displaced — they are redeployed from routine repeat questions to the complex cases the escalation protocol routes to them.

Response accuracy was maintained through mandatory source citation and real-time expert escalation for complex or clinically ambiguous cases. Volunteers were not displaced but redirected from routine repeat questions toward the cases the protocol escalates to them.

Figure 4. Patient-reported outcome domains

Three dimensions along which patients reported change, and the mechanism in each

Domain 01 — Emotional burden	Domain 02 — Medication adherence	Domain 03 — Access to specialist care
Reduced distress at the two points where it peaks — initial diagnosis and recurrence — when questions arise outside clinic hours and the alternative is unmoderated internet search.	Improved adherence through integrated dosing reminders and structured side-effect tracking, converting a one-off instruction into an ongoing prompt.	Accelerated specialist referral via AI-supported hospital and clinic navigation, shortening the interval between recognising a problem and reaching the right clinician.
Mechanism: immediate, plain-language explanation that acknowledges the fear before addressing the clinical content.	**Mechanism:** side-effect questions answered at the moment of doubt, rather than resolved by discontinuation.	**Mechanism:** navigation support addresses a system-literacy barrier, not a clinical one.

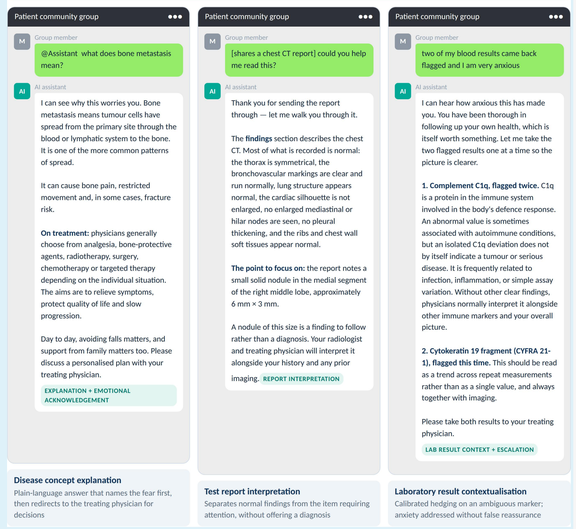
Patient-reported outcomes spanned reduced emotional burden at diagnosis and recurrence, improved medication adherence through dosing reminders and side-effect tracking, and accelerated specialist referral via hospital and clinic navigation.

Interaction content clustered into three recurring classes: explanation of disease concepts, interpretation of imaging and laboratory reports, and contextualisation of ambiguous results — each answered in plain language and returned to the treating physician for decisions.

Figure 5. Representative interaction types — de-identified reconstruction

Three recurring question classes observed in the deployment, reconstructed and translated from Chinese

De-identification notice These panels are reconstructions, not screenshots. All group member identifiers, which in the source system encoded diagnosis, stage and regimen, have been removed and replaced with generic labels; clinical details are representative rather than drawn from any individual patient. No patient data is reproduced.



The consistent pattern is that the assistant addresses the emotional register before the clinical content, and declines to convert an observation into a diagnosis. Escalation and physician referral are built into the answer rather than appended as a disclaimer.

Figure 6. From patient conversation to real-world evidence

The interaction log is a by-product of care delivery, not a separate data collection exercise

Step 01 — Patient question	Step 02 — Structured capture	Step 03 — Aggregation	Step 04 — Evidence asset
Asked in the patient's own words, unprompted by any instrument	Topic, disease stage context and concern type coded at scale	Patterns across thousands of sessions per month	Patient-reported concerns, unmet needs and treatment barriers

HEOR analysis	NRDL access argumentation	Patient-centred outcome measures
Patient-reported burden and treatment barriers captured continuously rather than at survey intervals	Documented unmet need and adherence barriers evidenced in patients' own terms	Instrument development grounded in what patients actually ask about

Timing advantage Signals surface months ahead of conventional clinical data. A patient asks about a side effect long before it appears in a pharmacovigilance dataset, and asks about cost long before non-persistence shows up in claims.

The structured interaction data generated a real-world evidence dataset of patient-reported disease concerns, unmet needs and treatment barriers — available months before conventional clinical data signals, because a patient asks about a problem long before it reaches a pharmacovigilance or claims dataset.

The deployment therefore yields two distinct assets from a single intervention: a scalable care delivery channel, and a continuous evidence stream. Conventional evaluation captures at most one of them.

Figure 7. Dual value proposition for rare disease value frameworks

The platform is simultaneously a care delivery innovation and an evidence generation mechanism

Value dimension	Evidence from deployment	Policy implication
Dimension 01 — Support scalability	~313 sessions/day sustained at zero incremental staffing cost, with 24/7 coverage unavailable under a volunteer-staffed model.	**Cost-effective care delivery** Marginal cost near zero per additional patient
Dimension 02 — Patient outcomes	Measurable change across emotional burden, medication adherence and speed of specialist referral.	**Health outcome contribution** Adherence and access are established value levers
Dimension 03 — Evidence generation	A continuous real-world dataset of patient-reported concerns, unmet needs and treatment barriers, available ahead of conventional signals.	**RWD infrastructure** Novel input to HEOR and access dossiers
Dimension 04 — Equity of reach	Support delivered independent of geography, reaching patients for whom specialist centres are a journey rather than a destination.	**Access equity** Value accrues where PAG infrastructure is thinnest

72,549 sessions in 232 days · **0** incremental staff cost — AI-enabled PAG platforms should enter rare disease value frameworks under both headings at once — as a care delivery innovation and as an RWD generation mechanism — since evaluating either in isolation understates the case.

Assessed under only one heading, the platform is undervalued by roughly the whole of the other.

Conclusions

AI chatbot deployment within rare disease patient communities demonstrates scalable, cost-effective patient support with measurable outcomes across emotional, adherence and access dimensions.

The real-world evidence generated by patient interactions provides a novel data source for HEOR analysis, NRDL access argumentation and patient-centred outcome measure development.

These findings suggest that AI-enabled PAG platforms should be incorporated into rare disease value frameworks as both a care delivery innovation and an RWD generation mechanism, with direct policy implications.

Reference

1. Panda and Friends patient group AI assistant: operating records, 1 January – 20 August 2025 (232 days; 72,549 sessions).
2. China Organisation for Rare Disorders (CORD) patient community partnership documentation, 2025.
3. 2025 Digital Life Physician (DLP) Survey. Base: n=4,500 Chinese physicians; AI users n=4,246.
4. Li A, Wang D, Zhou Y. Evolution of AI Value in Chinese Clinical Practice: 2025 Survey and Case Series. 2025.

Acknowledgement

The authors thank the patient community volunteers and the China Organisation for Rare Disorders for their partnership, and the Digital Life Physician research team for benchmarking context. All illustrative interactions shown are de-identified reconstructions; no patient data is reproduced.