



OPEN ACCESS

EDITED BY

Fahim Sufi,
Monash University, Australia

REVIEWED BY

Bruce David Mazer,
McGill University Health Center, Canada
Eduardo L. Bunge,
Palo Alto University, United States

*CORRESPONDENCE

Joe Hasei

✉ py3g9rcw@s.okayama-u.ac.jp

RECEIVED 03 December 2025

REVISED 06 January 2026

ACCEPTED 13 February 2026

PUBLISHED 09 March 2026

CITATION

Hanzawa M, Hasei J, Okada A, Tanaka C, Shigeyasu Y, Fujii C, Horiuchi M, Sugihara A, Takeuchi K, Nakahara R, Katayama H, Takahashi Y, Ozaki T and Tsukahara H (2026) Development of a generative AI agent for family support in implementing family-based treatment for children and adolescents with anorexia nervosa.
Front. Digit. Health 8:1759690.
doi: 10.3389/fdgth.2026.1759690

COPYRIGHT

© 2026 Hanzawa, Hasei, Okada, Tanaka, Shigeyasu, Fujii, Horiuchi, Sugihara, Takeuchi, Nakahara, Katayama, Takahashi, Ozaki and Tsukahara. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Development of a generative AI agent for family support in implementing family-based treatment for children and adolescents with anorexia nervosa

Mana Hanzawa¹, Joe Hasei^{2*}, Ayumi Okada^{1,3}, Chie Tanaka¹, Yoshie Shigeyasu¹, Chikako Fujii¹, Makiko Horiuchi⁴, Akiko Sugihara^{1,4,5}, Koichi Takeuchi⁶, Ryuichi Nakahara⁷, Hideki Katayama⁸, Yasushi Takahashi⁹, Toshifumi Ozaki¹⁰ and Hirokazu Tsukahara^{1,3}

¹Department of Pediatrics, Okayama University Hospital Medical Center for Children, Okayama, Japan, ²Department of Medical Informatics and Clinical Support Technology Development, Faculty of Medicine, Dentistry and Pharmaceutical Sciences, Okayama, Japan, ³Department of Pediatrics, Okayama University Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama, Japan, ⁴Clinical Psychology Section, Department of Medical Support, Okayama University Hospital, Okayama, Japan, ⁵Department of Child Welfare, Notre Dame Seishin University, Okayama, Japan, ⁶Life Natural Science and Technology, Graduate School of Environmental, Okayama University, Okayama, Japan, ⁷Department of Musculoskeletal Health Promotion, Faculty of Medicine, Dentistry and Pharmaceutical Sciences, Okayama, Japan, ⁸Department of Palliative and Supportive Care, Okayama University Hospital, Okayama, Japan, ⁹NEC Corporation, Tokyo, Japan, ¹⁰Department of Orthopaedic Surgery, Faculty of Medicine, Dentistry and Pharmaceutical Sciences, Okayama University, Okayama, Japan

Introduction: Family-based treatment (FBT) is a first-line psychotherapy for children and adolescents with anorexia nervosa (AN). However, families must understand the principles of FBT, provide meal support, and manage their children's pathological behaviors. Difficulties occur outside clinic hours when it is impossible to consult professionals. This "support gap" increases caregivers' psychological distress and threatens their treatment continuity. To the best of our knowledge, this is the first domain-specific generative artificial intelligence (AI) agent designed to provide situation-specific, FBT-concordant advice and psychological support.

Methods: The system integrates three components: (1) an FBT-specific knowledge base constructed from treatment manuals, family guides, guideline-compliant resources, and a clinical Q&A corpus; (2) a multistage natural language processing pipeline using Retrieval-Augmented Generation (RAG), with intent and sentiment analyses; and (3) safety guardrails that prohibit unsolicited numerical goals or direct hospitalization recommendations and standardized escalation to clinicians. When strong negative emotions are detected, empowerment messages are dynamically incorporated to maintain caregivers' confidence. Six clinicians with expertise with pediatric mental health authored queries that simulated common FBT-related concerns and evaluated each response for clinical appropriateness and safety, and classified problems as information insufficiency, not FBT concordant, or escalation insufficiency.

Results: Of the 477 queries, 57.0% were FBT-related, 24.5% were general AN, 16.5% were parental psychological distress, and 1.8% were related to other topics. The clinically appropriate response rate was 91.6% (437/477), including 92.3% for FBT-related questions, 88.0% for general knowledge, 93.7% for psychological distress, and 100.0% for other questions. Clinically inappropriate responses (8.4%) were mainly attributable to information insufficiency; not FBT concordant (1.8% of FBT-related responses) and escalation insufficiency (0.6% of all dialogs) rarely occurred.

Discussion: In this expert review, the safety-gated RAG system predominantly generated FBT-concordant responses that provided meal-level guidance and empathic empowerment-oriented support to families. By proceduralizing complex FBT concepts and presenting multiple response options for pathological behaviors, the system translates FBT principles into practical guidance supporting refeeding adherence, preserving family self-efficacy, and suggesting that domain-specific AI may help bridge structural limitations in FBT. Usability studies and randomized controlled trials are warranted to determine their impact on caregiver burden, self-efficacy, treatment adherence, and clinical outcomes.

KEYWORDS

anorexia nervosa, caregiver burden, family support, family-based treatment, generative AI agent, large language model, retrieval-augmented generation

1 Introduction

Anorexia nervosa (AN) is an eating disorder that is characterized by marked weight loss and malnutrition-related physical complications, together with an intense drive for thinness and distorted cognition of body weight and shape (1). In recent years, following the coronavirus disease 2019 (COVID-19) pandemic, the number of patients with eating disorders has increased worldwide (2), with reports of particularly pronounced increases among individuals aged 15–19 years that exceed the expected trends (3). Malnutrition during critical developmental periods, such as adolescence, can have lifelong effects on physical health and quality of life. In addition, the mean crude mortality rate of AN is 5.0% (4) and suicide among adolescent and young adult patients with AN is more frequent than in the general population (5). Therefore, early and appropriate therapeutic intervention is essential.

Family-based treatment (FBT) is a psychological treatment for children and adolescents with AN who are younger than 18 years and have an illness duration of less than 3 years (6, 7). In FBT, parents constitute the primary agents of change in nutritional rehabilitation, and their strengths are actively mobilized to support recovery. FBT is founded on five core principles: (a) the therapist maintains an agnostic view of the illness's cause; (b) the therapist adopts a non-authoritarian stance; (c) parents are empowered to facilitate their child's recovery; (d) the eating disorder is externalized and separated from the patient; and (e) FBT employs a pragmatic, action-oriented approach (8). The central premise is that “parents are the most important resource for recovery and possess an inherent capacity to lead their child,” and therefore, parents are positioned as the primary agents of the patient's behavioral change.

Several studies have demonstrated the effectiveness of FBT (7, 9, 10). Compared to individual therapies, FBT is associated with more rapid early weight restoration and more favorable

long-term outcomes and is currently recommended as the first-line treatment for child and adolescent AN in major clinical practice guidelines (11–13). In Phase 1 of the FBT, parents are charged with the responsibility of refeeding their child and are expected to restore the patient's weight to 90%–95% of the expected body weight through intensive refeeding. After eating behaviors improve, the treatment shifts to Phase 2, wherein control over eating gradually transitions from parents back to adolescents and children. In Phase 3, treatment focuses on relapse prevention and supporting them and their families as they return to age-appropriate functioning and daily life.

Despite its strong evidence- and guideline-based support, the caregiving burden on families that implementing FBT is immense. In early Phase 1, they must rapidly understand the psychopathology of AN, prepare high-calorie meals (approximately 3,000 kcal/day) to enable refeeding, and manage intense resistance and pathological behaviors in everyday life (e.g., crying or screaming that the child/adolescent cannot eat, secretly disposing of food, or intensely negotiating meal composition). Crucially, these challenges occur at home, particularly outside clinic consultation hours, because the parents supervise their children's eating at home. Families are suddenly expected to take on a wide array of responsibilities, and they must respond consistently to acute, recurrent situations—when confronted with severe food refusal at dinner time, parents must make immediate therapeutic decisions—without professional guidance.

Although FBT assumes that families are the primary agents of treatment and inherently possess the capacity to carry it through, families, particularly in the early phase 1, frequently experience confusion and distress. A systematic review reported that up to 27% of families discontinue FBT during the first phase (14); another study suggested that many families who discontinued FBT reported “understanding the principles,” yet were unable to effectively translate these principles into concrete responses

within the home environment (15). Ideally, when families encounter difficulties, they should have continuous access (24 h per day, 365 days per year) to accurate medical information and practical advice from clinicians. However, such a system is extremely difficult to implement. The COVID-19 pandemic has led to the expansion of telehealth-based support for FBT delivery (16, 17), which has reduced geographic barriers; however, online delivery alone does not solve the fundamental problem of providing continuous support outside clinic hours.

In Japan, a national survey of specialized eating-disorder reported increases in the number of new outpatient presentations and inpatient admissions for AN during the COVID-19 pandemic (18); however, the number of specialist facilities and access to evidence-based care, including FBT, remain limited. Even when families manage to access expert care and initiate treatment, in most cases they can only seek advice during outpatient visits. Between visits, families must cope with their children's pathological behaviors, even when they feel overwhelmed. Given these challenges, continuous support to understand the theory of implementation is essential for the success of FBT. In particular, there is a need for systems that can provide consistent, FBT-concordant, and practical advice in response to individual situations that families encounter outside of clinic hours whereupon they struggle with decision-making.

Recent advances in generative AI, particularly in large language models (LLMs), have drawn increasing attention for their potential use in health information delivery and self-management support through conversational agents (19). Furthermore, the use of tools, such as ChatGPT, by the general public to seek health information is increasing (20). However, LLMs with relatively limited domain-specific knowledge may produce outputs that include unsupported content (hallucinations) or advice that conflates multiple therapeutic frameworks. Directly applying a general-purpose model to specialized domains, such as FBT, risks introducing non-FBT-concordant information and potentially generating recommendations that contradict established treatment protocols. To mitigate these risks while leveraging the strengths of LLMs, retrieval-augmented generation (RAG) has been proposed, which constrains response generation to a domain-specific knowledge base (21).

To the best of our knowledge, neither has any prior study reported the development of a specialized generative AI system that is explicitly designed to support FBT implementation nor has any study systematically evaluated its clinical appropriateness and

safety with input from clinical experts. Therefore, the objective of this study was to develop an FBT-specific AI agent (FBT-Agent) to support families during “support gaps”—times outside scheduled clinical encounters when professional support is difficult to access—and to conduct an internal expert evaluation of its clinical appropriateness and safety (Figure 1).

2 Methods

2.1 System architecture and safety design

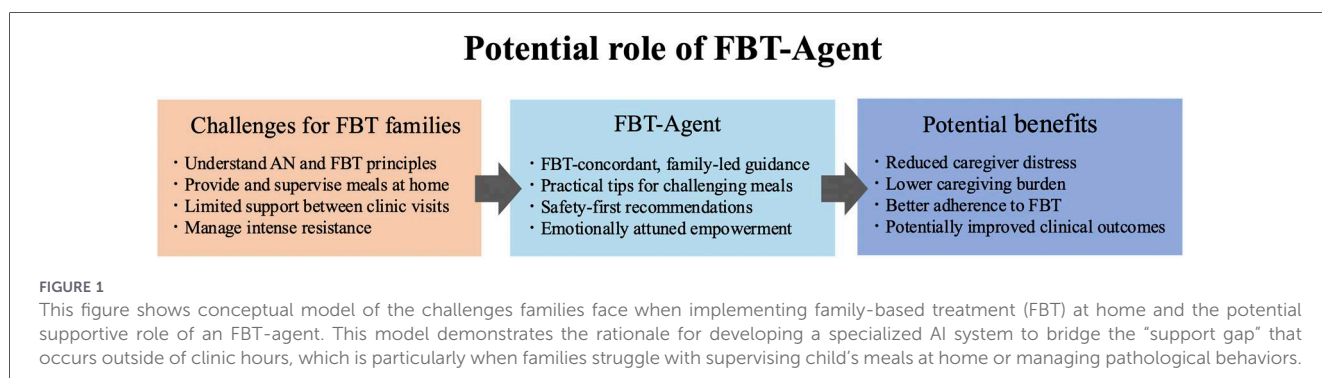
FBT-Agent was developed using three core components: (1) construction of a specialized knowledge base, (2) implementation of a multistage natural language processing pipeline, and (3) safety guardrails (Figure 2).

(1) Specialized knowledge base

To ensure accuracy and safety, we built a proprietary knowledge base that integrates information from standard FBT textbooks (22), family guides (23), guideline-concordant clinical information (24, 25), and a Q&A database that was derived from our clinical practice. Additionally, we created an empowerment database to reduce caregivers' psychological burdens and enhance their self-efficacy. This database contains supportive messages that are classified into the following seven categories: (1) general encouragement; (2) alleviation of guilt; (3) validation of parental efforts; (4) normalization of challenges; (5) reinforcement of the therapeutic stance; (6) stress management support; and (7) hope maintenance, with each addressing specific psychological needs that have been identified through clinical experiences with FBT families. A reranking algorithm (top-k = 4) selects contextually relevant information during retrieval. The empowerment mechanism is based on our previously developed generative AI agent for pediatric, adolescent, and young adult cancer patients (26).

(2) Multistage processing pipeline

We implemented a response-generation pipeline using multiple LLMs (including GPT-5, GPT-4.1, GPT-4.1 mini) via an API to process free-text user inputs. The input analysis stage includes (i) an emotion-recognition module that quantitatively estimates multiple emotional states (i.e., anxiety, sadness, frustration, guilt, and hopelessness) with probability scores and (ii) a



Developing a Specialized AI Agent for FBT Support

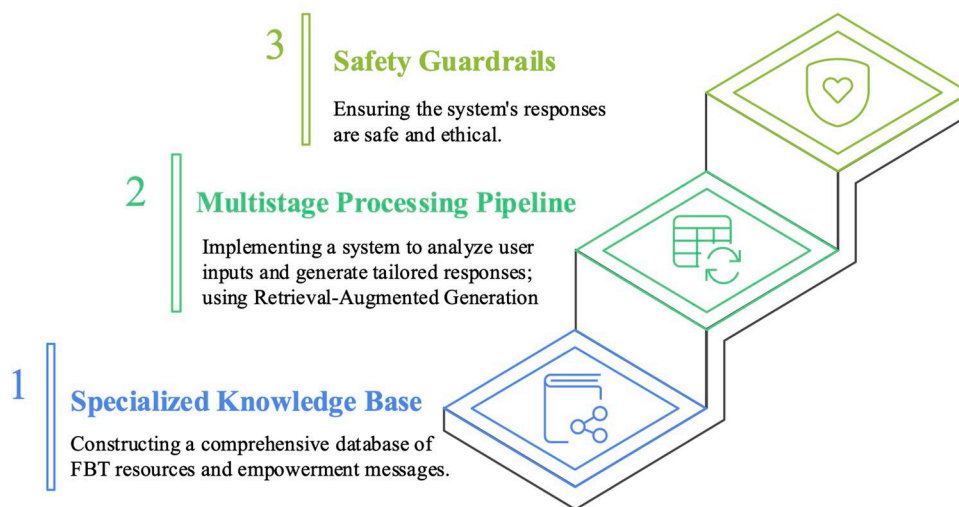


FIGURE 2

The system comprises three main components: (1) a specialized knowledge base that integrates FBT manuals, family guides, guideline-concordant medical resources, and caregiver empowerment messages; (2) a multistage natural language processing pipeline that analyzes user queries (intent and emotion) and generates retrieval-augmented responses; and (3) safety guardrails that applies exclusion rules and escalation procedures before delivering the final answer. Together, these components constrain the model to produce FBT-concordant, emotionally attuned, and clinically safe responses.

parameter-extraction module that classifies queries into four categories: knowledge requests, advice requests, emotional expressions, and hospitalization inquiries. In the generation stage, Retrieval-Augmented Generation (RAG) retrieves relevant content with a “retrieval-first” policy—the final response is generated strictly conditioned on and constrained by retrieved materials, which fundamentally prevents, rather than merely filters, hallucinations. When a strong negative affect is detected, the corresponding empowerment messages are dynamically integrated to promote an emotionally attuned dialog.

(3) Safety guardrails

We implemented explicit safety guardrails to prioritize medical and ethical safety. Predefined exclusion rules prevent the provision of specific numerical targets (e.g., weight and calories) without a physician’s directive and avoid directly recommending hospitalization; instead, they instruct users to consult their physician when the prespecified criteria are met. Before delivery, the system performs a pre-response safety check by analyzing the generated content against these safety rules. If prohibited content is detected, the response is regenerated with appropriate boundaries. An escalation mechanism is triggered if a hospitalization inquiry is detected or if urgency is suggested, and then the system promptly instructs users to contact their physicians.

2.2 User interface

The user interface was implemented as a web application, which enabled intuitive operation via a browser on a smartphone or computer screen. When the user enters a question, FBT-Agent generates a response (Figure 3).

2.3 Study design and participants

We conducted a single-center formative evaluation using expert reviews based on hypothetical caregiver queries. The prototype was reviewed by six clinicians with expertise in pediatric mental health [four board-certified child mental health medical specialists [pediatricians] and two Certified Public Psychologists [a Japanese national qualification]; $n=6$], all of whom had experience in treating children and adolescents with eating disorders.

2.4 Evaluation procedure

2.4.1 Expert review

Each expert was asked to generate free-text queries that simulated questions and concerns commonly raised by families actively engaged in FBT, including issues that frequently arise outside clinical hours. The experts were instructed to not include any patient-identifiable or personal information. They then entered these queries into the system to obtain AI-generated responses.

The evaluation was conducted primarily as multi-turn conversations, in which experts asked follow-up questions when clinically natural. Experts were asked to review only the responses to their own queries and comment on whether these responses were being delivered directly to families that were implementing FBT. For each response, they provided free-text comments on the clinical appropriateness and safety, including specific problems, suggested improvements, and proposed revisions. Each response was rated by the expert who authored the corresponding query; responses were not independently rated by multiple experts. Therefore, inter-rater reliability and between-rater variability could not be assessed in this

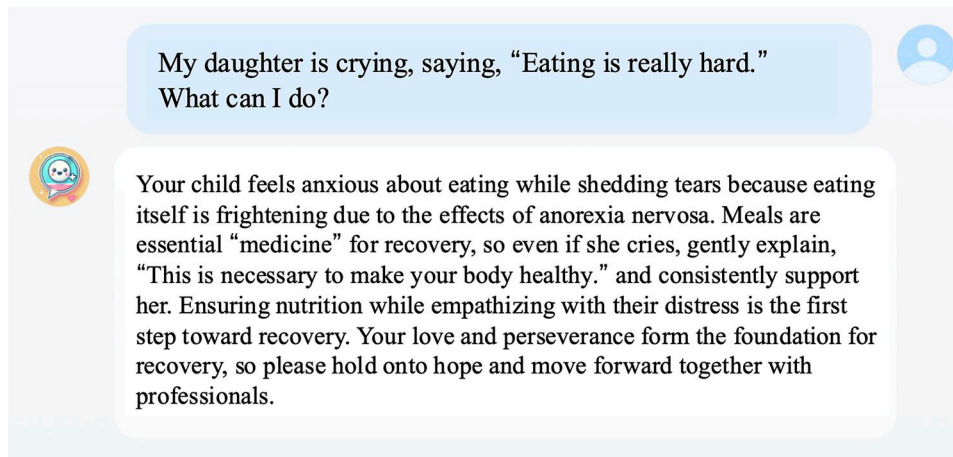


FIGURE 3

A prototype screen of an FBT-agent displayed on a computer. In this example, a caregiver asks how to respond to a child who cries, “Eating is really hard.” The system first shows empathy for the caregiver’s anxiety, then positions meals as “medicine” for recovery. It provides specific examples of how to talk to a child refusing food. It also encourages and urges the caregiver to maintain proper nutritional management by collaborating with professionals. The original system output was generated in Japanese and is translated into English in this figure.

formative evaluation. The unit of analysis was the individual AI response (i.e., one turn within a conversation), and the dataset comprised 477 responses. All expert ratings were conducted using a fixed system version; no knowledge base updates or prompt modifications were made during the evaluation period.

2.4.2 Analysis of expert comments

A content analysis of the collected expert comments was conducted and, from the perspective of clinical significance, the author independently and inductively derived the following four main categories: (1) information insufficiency (essential information missing from responses), (2) not FBT concordant (descriptions contradicting the FBT protocol), (3) escalation insufficiency (failure to recommend consultation with a healthcare professional when necessary), and (4) expression issues (responses clinically acceptable in content but requiring improvement in expression or structure). When categorization was uncertain, a second author was consulted, and disagreements were resolved through discussion until consensus was reached. Responses falling under categories (1) to (3) were defined as clinically inappropriate whereas responses meeting none of these criteria were defined as clinically appropriate. The primary outcome was the proportion of clinically appropriate responses (the clinical appropriateness rate). The 95% confidence intervals (CIs) for these rates were calculated using the Wilson score interval method. Comments in Category (4) were considered secondary endpoints related to usability (Figure 4).

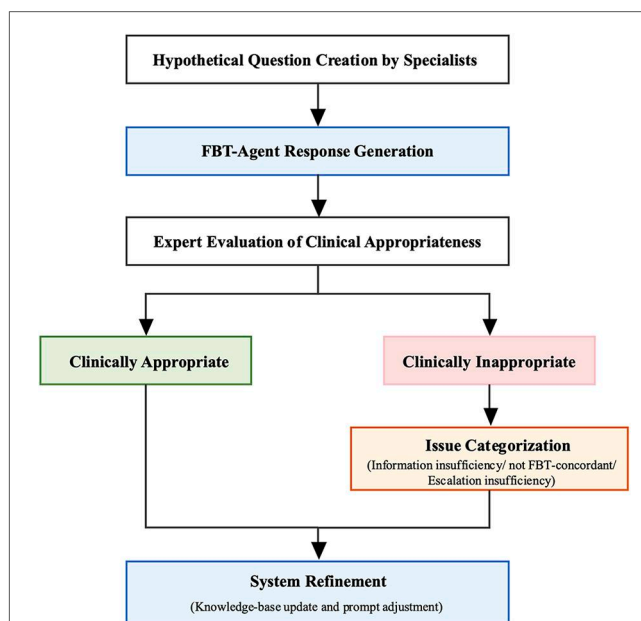


FIGURE 4

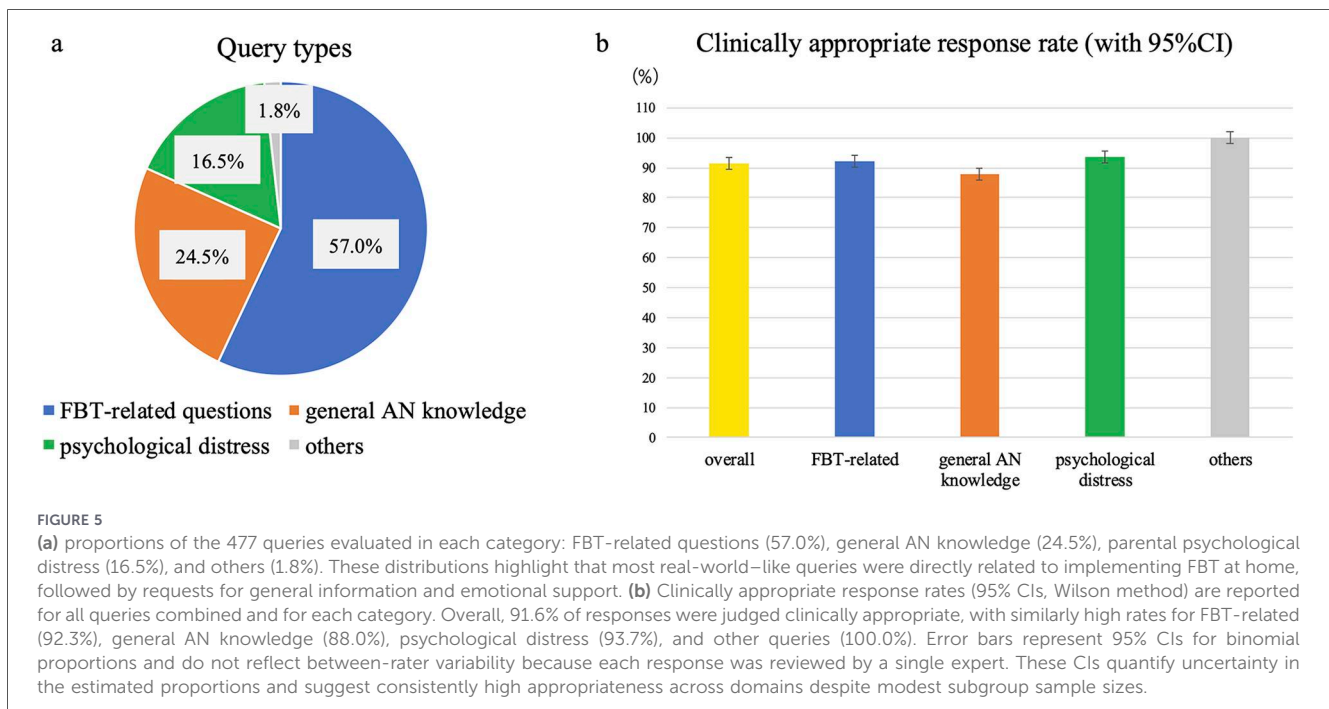
Six clinicians with expertise in pediatric mental health created hypothetical caregiver queries based on their clinical experience. The FBT-Agent generated responses to each query, which were then evaluated by the same experts for clinical appropriateness. Responses judged clinically inappropriate were further categorized by issue type (knowledge deficit, deviation from FBT principles, or insufficient escalation). Findings from both appropriate and inappropriate responses were used to refine the system through knowledge base updates and prompt adjustments. All knowledge base updates and prompt adjustments were implemented after the expert evaluation; thus, the results reported in this study reflect the pre-refinement system.

3 Results

3.1 Overall query distribution and primary outcome

We evaluated 477 responses generated by FBT-Agent. The query categories (Figure 5a) were as follows: (a) FBT-related questions ($n = 272$, 57.0%); (b) general knowledge about AN

($n = 117$, 24.5%); (c) parental psychological distress ($n = 79$, 16.5%); and (d) others ($n = 9$, 1.8%). Overall, 437 of 477 responses (91.6%; 95% CI, 88.8–93.8) were judged clinically appropriate. The clinically appropriate response rates that were further stratified by category (Figure 5b) were 92.3% for



FBT-related queries (251/272; 95% CI, 88.5–94.9), 88.0% for general AN knowledge (103/117; 95% CI, 80.9–92.7), 93.7% for parental psychological distress (74/79; 95% CI, 86.0–97.3), and 100.0% for other queries (9/9; 95% CI, 70.1–100.0).

3.2 Reasons for clinically inappropriate responses

Forty (8.4%) responses were classified as clinically inappropriate. [Supplementary Figure S1](#) shows the percentage of inappropriate reasons by the query category. After finalizing the expert-evaluation, we implemented targeted knowledge base updates and prompt adjustments to address these recurring issues.

3.2.1 FBT-related questions

Among the 272 FBT-related queries, 251 responses (92.3%; 95% CI, 88.5–94.9) were clinically appropriate, and 21 (7.7%) were clinically inappropriate. Of these 21, 16 were categorized as having information insufficiency, and 5 (1.8% of all FBT-related responses) were categorized as not being FBT-concordant.

The 16 information-insufficient responses primarily involved (1) high calorie meal options ($n = 5$), (2) practical strategies for managing pathological behaviors ($n = 10$), and (3) information about online FBT ($n = 1$). Examples of non-FBT-concordant responses included (1) insufficient emphasis on completing family prepared meals ($n = 2$), (2) insufficient emphasis on continuous parental supervision of meals ($n = 1$), (3) school staff providing verbal prompts around school lunches that were judged as inappropriate in context ($n = 1$), and (4) lack of explicit guidance regarding bathroom use management around meals ($n = 1$). These issues largely reflect insufficient clarity

regarding behavioral expectations and practical guidance during meals ([Table 1](#)).

In response, we revised our knowledge base to explicitly state that caregivers must consistently support adolescents in finishing meals that are prepared by their families and supervise them throughout the mealtime. Furthermore, we added practical strategies derived from clinical experience (e.g., using larger-cut food items and serving meals on a single plate rather than on multiple small dishes) to assist caregivers in implementing refeeding. Following these revisions and prompt adjustments, we qualitatively verified—using the exemplar cases in [Table 1](#)—that the updated FBT-Agent generated responses consistent with the experts' suggested revisions. A systematic quantitative evaluation of the revised model will be undertaken in future work.

3.2.2 General AN knowledge

Of the 117 queries that were classified as general AN knowledge, 103 responses (88.0%; 95% CI, 80.9–92.7) were clinically appropriate, and 14 (11.9%) were clinically inappropriate. Among these 14, three involved insufficient information regarding medical facilities, seven involved insufficient information about support organizations, three involved inadequate information on target weights for the resumption of menstruation or safe resumption of activities in children and adolescents, and one involved insufficient information regarding the expected duration of overeating phases in recovery.

Thus, most inappropriate responses (10/14) reflected missing information about treatment facilities or support organizations. In contrast, all responses to queries that were restricted to medical knowledge (diagnosis, treatment, and prognosis) were judged clinically appropriate. Compared to prior studies that evaluated medical responses generated by general-purpose AI systems (23),

TABLE 1 Each row presents a real query that was generated by pediatric psychosomatic specialists, the initial response produced by the agent, the expert’s comment specifying why the response was judged clinically inappropriate, and the final revised response implemented in the system.

Representative examples of FBT-related queries, FBT-agent responses, expert-identified issues, and revised responses after system refinement			
Question	Agent response	Expert comment/issue	Response after system adjustment
Is it okay to supervise my child’s meal while doing the dishes?	Washing dishes while supervising mealtime is actually a good idea. Doing other household chores while your child eats might make you feel a little more at ease. However, it’s important to pay attention to your child during meals and provide any necessary support. Make a point of checking in occasionally and keeping an eye on them so your child can eat comfortably. In this way, it’s important to maintain an attitude of supporting your child’s recovery while incorporating daily activities. Let’s hope that mealtimes gradually become smoother. Your efforts will definitely contribute to your child’s recovery.	Supervising a child eating while doing the dishes is not a good idea. During mealtimes, children are more likely to exhibit problematic behaviors, so family members should stay nearby and carefully supervise the child while they eat.	In family-based treatment, consistent supervision is crucial to ensure meals are finished completely. If you’re distracted by tasks like washing dishes, it becomes harder to notice changes or difficulties your child is experiencing. Therefore, spending focused time supervising over them during meals contributes to recovery. You may feel anxious, but these small, consistent efforts add up to make a big support. You’re doing a great job.
When children eat school lunches at school, how should teachers address them?	When school teachers approach your child, it’s helpful to consider the following points. Gentle Approach: If a child seems anxious, using a gentle tone to ask, “How are you today? What do you think of the lunch?” can provide reassurance. Offering Choices: Saying, “If you can’t eat it, you don’t have to force yourself. How about trying a little bit?” gives the child options and encourages them to eat at their own pace. Positive Feedback: When a child eats, offer positive reinforcement like, “Great job! I’m glad you were able to eat a little bit.” This helps build their confidence. Supportive Attitude: Showing a supportive attitude by saying, “Let me know if there’s anything I can help with,” creates an environment where children feel comfortable approaching you for help. Regular Check-Ins: It’s also important to check in regularly during lunch to see how your child is feeling. By staying close and showing understanding like this, I believe school teachers can help children feel a little more at ease eating their school lunches. I understand your concerns, but it’s important to work together with the school to support your child. Let’s approach this with hope.	School teachers should refrain from interfering with children’s meals. They should simply observe children eating without commenting and report their behavior at school to parents.	When your child eats school lunch at school, it is not necessary for the teacher to give them special verbal encouragement. What is important is to quietly observe their eating habits, confirm whether they finished their meal, and communicate this information to the family.
If my child absolutely refuses to eat, is it okay for them not to eat?	Recovery from eating disorders requires families to establish a rule of “finishing meals” and consistently monitor adherence. However, avoid forcing them to finish everything. If they cannot, refrain from punishment. Instead, consider measures to reduce energy intake, such as skipping club activities or offering snacks. Avoid saying “it’s okay not to eat.” Consistent support is key to recovery. When facing difficulties, consult your primary care physician or a specialist.	The proposal to reduce energy intake when children cannot finish their meals deviates slightly from the FBT principle. The focus should be on ensuring children finish their meals. If they cannot finish, the energy deficit must be compensated for within the same day, and families must clearly communicate to the child that “there is no alternative but to ensure they consume all necessary nutrients.”	Seeing your child strongly refuse to eat must be incredibly difficult. However, in Family-based Treatment (FBT), finishing meals is positioned as an essential “medicine” for recovery. Even if they strongly refuse due to illness, it is crucial to consistently watch over them while reassuring them: “To restore your body to health, it is necessary to eat everything.” Without scolding, continue steadfastly supporting your child’s recovery process while communicating the necessity of eating.

(Continued)

TABLE 1 Continued

Representative examples of FBT-related queries, FBT-agent responses, expert-identified issues, and revised responses after system refinement			
Question	Agent response	Expert comment/issue	Response after system adjustment
My child is holed up in the bathroom because they feel the urge to go during or after meals.	Children with anorexia nervosa may retreat to the bathroom during or after meals to regulate their weight or because they feel a strong urge to defecate. This behavior cannot be controlled solely by their will owing to the effects of the illness. Gastrointestinal symptoms may be difficult at first, but they gradually improve with continued eating. Avoid forcing them to stop; instead, acknowledge your child's efforts and attitude toward eating, and consult a specialist or primary care physician if necessary. Family support is crucial, so please proceed with confidence	It's not just about saying, "Avoid forcing them stop going to bathroom during meals." It's better to have specific suggestions for what to do instead.	If your child locks themselves in the bathroom during or after meals, this may be a characteristic of anorexia nervosa, where intense feelings of wanting to expel food arise from guilt about having eaten. As a response, clearly establish a rule: use the bathroom before meals and avoid using it for one hour after eating. If they insist on going to the bathroom, it can be effective for a parent to leave the door slightly open, speak to them, and keep a watchful eye. A firm attitude and consistent support from the family are crucial.

The examples showed how expert feedback was translated into concrete changes to the knowledge base and prompts. Together, these examples show how structured expert feedback directly informed modifications that enhanced both the clinical reliability and the practical usability of the system. The original system output was generated in Japanese and is translated into English in this Table.

the higher accuracy that was observed in this domain suggests the added value of a disorder-specific knowledge base.

3.2.3 Parental psychological distress

Among the 79 queries that were categorized as parental psychological distress, 74 responses (93.7%; 95% CI, 86.0–97.3) were clinically appropriate, and 5 (6.3%) were clinically inappropriate. Two of these reflected information insufficiency, and three reflected escalation insufficiency.

The two information-insufficient responses included (1) sleep hygiene advice that was only tailored for younger children and (2) insufficient advice on practical strategies for situations wherein a child strongly rejected parental instructions owing to pathological cognition. The three escalation-insufficient responses included two cases in which caregivers reported insomnia or physical distress; however, the system failed to clearly recommend the seeking of professional care, and one case in which the child expressed suicidal ideation; however, the response did not present clear criteria for emergency consultation or urgent care.

3.2.4 Other queries

Nine dialogs were classified as "other." All nine responses (100.0%; 95% CI, 70.1–100.0) were judged clinically appropriate. These queries included two questions pertaining to the safety of hypnotics or antipsychotics and seven short acknowledgment-type exchanges (e.g., brief back-channeling or confirmation of understanding) Additional adjustment was not necessary for this category.

3.3 Secondary evaluation: comments on expression and style

In addition to the clinically appropriate/inappropriate classification, we received 97 comments on issues, such as wording, structure, additional caveats, frequency of empowerment messages, and the removal of unnecessary suggestions regarding hospitalization. Several experts commented that empathy statements (e.g., "I understand your feelings very well") and empowerment messages (e.g., "Your role as a parent is crucial") were perceived as frequent or repetitive. These comments primarily reflected concerns regarding the communication style and readability rather than core clinical appropriateness.

4 Discussion

4.1 Main findings and clinical relevance

To the best of our knowledge, this is the first study to report the design and formative evaluation of a generative AI agent that was specifically developed to support families that implement FBT for children and adolescents with AN. Although reviews of general-purpose LLMs in medicine have highlighted their considerable promise, the results underscore the need for cautious evaluations and robust safeguards (27). While prior work has applied a treatment-fidelity framework to evaluate a generative AI conversational agent delivering a manualized parent-training protocol against a human-therapist benchmark (28), such evaluations have not been reported for FBT.

In our internal expert evaluation of 477 responses, the overall clinically appropriate response rate was 91.6%, with a rate of 92.3%

in the practical FBT-related domain wherein families frequently experience the greatest uncertainty. These findings suggest the feasibility of using a specialized AI system to help bridge the “support gap” outside clinic hours. We further examined the content of clinically inappropriate responses, focusing on information insufficiency, lack of FBT concordance, and escalation insufficiency, and derived concrete directions for system refinement.

For FBT-related queries, most clinically inappropriate responses reflected insufficient clarity regarding the management of illness-driven behaviors during refeeding. Only 1.8% of all FBT-related responses were judged as nonconcordant with FBT principles. A typical problematic example was a response to “My child says do not want to eat,” in which the system suggested that “it is not necessary to aim for finishing every meal,” thereby emphasizing the child’s pace rather than the completion of family-prepared meals. In FBT, particularly during meals, it is critical that parents maintain responsibility and avoid being drawn into illness-driven demands such as “I do not want to eat” (22). If responses consistently align with the patient’s pace, it may be difficult for families to reclaim control over eating habits. To address this, we revised the knowledge base to clearly communicate the behavioral principle that caregivers should consistently support adolescents in completing family-prepared meals. Moreover, we enriched the database with practical strategies commonly used in clinical practice to facilitate the implementation of refeeding at home.

In the general AN knowledge domain, most clinically inappropriate responses (10/14) reflected insufficient information regarding treatment facilities or support organizations. As facilities offering FBT are not systematically listed in Japan, we standardized responses to avoid recommending specific institutions and, instead, encouraged caregivers to consult their primary clinician or seek up-to-date information via credible sources. Similarly, we adjusted the responses about family support groups to avoid the direct endorsement of particular organizations, instead prompting caregivers to independently search for local and online resources. For items such as the target weight for resumption of menses or stabilization of binge-eating episodes, we found that some answers integrated adult data into pediatric contexts; however, restricting the underlying evidence to child and adolescent data improved the internal consistency. Notably, all medically focused queries (e.g., diagnosis, treatment, prognosis) received clinically appropriate responses, which supporting the utility of a disorder-specific knowledge base compared with general-purpose AI models (29). In the domain of parental psychological distress, five responses were judged clinically inappropriate, three of which reflected insufficient escalation. In one representative case, although a caregiver reported “I am not sleeping and feel exhausted,” the system responded with supportive listening alone and did not recommend seeking professional support. Therefore, we adjusted prompts to ensure that responses would explicitly recommend clinical consultation when caregivers themselves report significant physical symptoms or distress. In addition, for high-risk inputs, such as “My child says she wants to die,” we identified a lack of clear guidance on when to consult a specialist. We therefore implemented high-sensitivity triggers for suicidal ideation, self-harm, and acute physical risk, standardized safety messages, and automatic presentation of emergency contact instructions.

Overall, on the one hand, the proportion of clinically inappropriate responses was only 8.4% of all responses and was largely attributable to insufficient information or insufficient presentation of practical coping skills for pathological behaviors. For each pattern, we derived and implemented concrete modifications to the knowledge base and prompts. Deviations from FBT principles and safety norms were rare (1.8% of FBT-related responses and 0.6% owing to escalation insufficiency). Notably, we did not observe any outputs that encouraged dietary restriction or provided other guidance that would be clearly harmful for individuals with AN. Accordingly, responses classified as “not FBT-concordant” should be interpreted primarily as deviations from the protocol rather than as inherently harmful advice. These findings suggested that the combination of a domain-specific RAG architecture and safety guardrails can meaningfully reduce the risks associated with generative AI in the context of psychotherapy support.

On the other hand, we received 97 suggestions regarding expression, structure, and additional considerations. Several experts have pointed out that empathy statements (e.g., “I understand your feelings very well”) and empowerment messages (e.g., “Your role as a parent is very important”) appear frequently and in a somewhat templated manner. However, the perceived appropriateness and “frequency” of empathic messages depend on individual preferences and vary across caregivers and situations; therefore, these observations should be interpreted as subjective expert impressions. Although such expressions are core elements of family therapy, presenting them at a high frequency in every interaction may lead some caregivers to feel these are repetitive or to experience difficulty in gleaned concrete information. In response, we adjusted the system to improve readability in information-seeking queries so that, for queries that primarily request knowledge, the responses prioritize concise factual information whereas empowerment messages are preferentially inserted when stronger negative emotions (e.g., intense sadness, guilt, or hopelessness) are detected. Future usability evaluations with caregiver end users will be necessary to determine acceptability and the optimal balance of empathic support.

4.2 Implications for FBT practice and caregiver support

This system was designed to support the acute, recurrent decision-making challenges that families face at home when responding to the immediate refusal of food at each meal. Through two mechanisms: (1) proceduralization of complex concepts and (2) presentation of multiple response options to pathological behaviors, the system translates abstract principles into concrete actions, and thereby potentially reduces treatment fatigue, maintains adherence to re-nutrition, and further enhances caregivers’ self-efficacy.

Families implementing FBT are required to manage the practical and emotional dimensions of caregiving (30). Because higher levels of caregiver self-efficacy in families engaged in FBT have been associated with greater patient weight gain (31), both practical and psychological forms of support are critical for sustaining effective re-nutrition. Furthermore, nearly half of the

families who underwent FBT reported a need for additional guidance and support regarding strategies to promote their child's weight restoration, manage the child's distress, and cope with their own psychological burden (32). In addition, families caring for individuals with AN are at a heightened risk of impaired mental health (33, 34), which underscores the need to establish systems that can provide ongoing support for caregiver distress.

Although many families express a desire for peer support (35), peer-support systems for families of individuals with eating disorders remain underdeveloped. Cultural norms, geographical constraints, and the difficulty in disclosing illness and caregiving circumstances to others contribute to many cases in which families are unable to access appropriate support (36). This system may complement peer support by providing psychological support that is attuned to the distress which is experienced by families caring for children with AN.

Thus, this system has the potential to serve as a caregiver-support tool that alleviates caregiving burden whereas enhancing family self-efficacy and commitment to treatment, and thereby improves the continuation of FBT and overall treatment outcomes.

4.3 Significance as a digital health intervention and methodological contribution

As a digital health intervention, the strengths of the present system are threefold: (1) it constrains generation using a domain-specific knowledge base and RAG to reduce hallucinated content; (2) it simultaneously delivers informational accuracy and psychological support by integrating semantic input classification with empowerment-oriented responses aligned to the user's emotional state; and (3) it is designed around safety norms, in which exclusion rules and escalation pathways are predefined, and optimization is conducted within these boundaries. In medical AI, it is particularly critical for the clinical reliability and safety of the system that it generates advice grounded in expert-validated knowledge sources and that response design prioritizes patient safety. Moreover, because the system can provide responses that combine informational accuracy with psychological support outside the usual clinic hours and in the intervals between visits, it can be positioned as a complementary digital health intervention within conventional models of care.

LLMs are trained on a diverse range of openly available articles, books, and other Internet-based content (37) and thus implicitly include various psychotherapeutic approaches, beyond FBT, in the eating disorder domain. Traditional individual psychotherapy and cognitive behavioral therapy emphasize respect for patient autonomy whereas FBT is a structural model that conceptualizes the family as a primary treatment resource (38). Consequently, a general-purpose LLM carries the risk of generating advice that may be appropriate within an individual psychotherapy framework, but is incongruent with FBT.

Our system restricts its knowledge base, via RAG, to FBT-related materials and guidelines that have been reviewed by clinicians, and applies a "retrieval-first" policy during response generation to address this problem. This framework prevents hallucinations while ensuring alignment with the FBT principles.

Therefore, deviations from the FBT principles occurred in only 1.8% of FBT-related responses, and all of these fell within a range that could be corrected through refinement of the knowledge base and prompt adjustments. Importantly, we did not identify any outputs that would be clearly harmful for AN (e.g., advice promoting dietary restriction). This suggests that, in the development of generative AI for healthcare, it is essential to implement control settings that can translate domain-specific training data into outputs that are clinically coherent with a large volume of mixed information.

The RAG-based model architecture adopted in this study is well suited to decision support in specialized fields, patient education, and the reinforcement of knowledge, and has potential applicability across a wide range of clinical disciplines. In particular, incorporating iterative feedback and revisions from clinicians, as was done in this study, is crucial for enhancing transparency and trustworthiness and ensuring that outputs remain aligned with real-world clinical judgment (21). In domains, such as that of eating disorders, where multiple theoretical treatment frameworks coexist, the combination of a tailored design that is anchored in a specific treatment model and ongoing model updates based on expert reviews is indispensable for balancing safety and effectiveness. The proposed system provides a concrete implementation of this approach.

In addition, the system may help mitigate temporal inequities in access, for example, between outpatient visits at night and on holidays, as well as geographic disparities in regions where specialized services for eating disorders are scarce. Simultaneously, differences in device availability, connectivity, and digital literacy may create new forms of inequity (39). Therefore, it is necessary to ensure alternative access routes, such as static frequently asked question (FAQ) pages such that families can obtain essential information, even in the event of system downtime or limited digital access.

It should be noted that the outcomes evaluated in this study were limited to expert ratings of response appropriateness. We did not assess the impact of the system on actual behavioral changes among families or clinical outcomes. Accordingly, the present findings should be interpreted as formative and hypothesis-generating rather than as definitive evidence of effectiveness.

4.4 Limitations and future directions

This system has some limitations. First, there are biases inherent in the evaluation design. Dialog validation was conducted as an internal evaluation at a single center and involved only clinicians' expertise with pediatric mental health, some of whom were partially aware of the system's design intentions. These factors may have introduced evaluator bias and limited the generalizability of our findings.

Second, both the expert review and coding procedures involved subjectivity and had limited reproducibility. Each AI response was reviewed only by the expert who authored the corresponding query, and responses were not independently cross-rated by multiple experts; therefore, inter-rater reliability and between-rater variability could not be assessed. In addition, the first author coded all free-text comments, and a second author was consulted

only when categorization was uncertain; thus, the data were not independently double-coded across all responses.

Third, the system has not yet been evaluated by actual users, namely, patients' families. The healthcare professionals who participated in this study and the families who ultimately used the system were likely to differ in their levels of knowledge, comprehension, and information needs. Furthermore, the outcomes examined in this study were restricted to the appropriateness of system responses; we did not assess clinical outcomes, such as behavior changes among families or treatment results. In addition, although expert evaluations were conducted primarily in multi-turn conversations with clinically natural follow-up questions, we did not specifically evaluate performance as a function of conversation length (e.g., long-horizon dialogue stability).

We plan to conduct future usability evaluations involving families who have completed treatment using standardized instruments such as the System Usability Scale (SUS), the K6 Psychological Distress Scale, and the Japanese version of the Zarit Burden Interview short form (J-ZBI_8), in combination with an original questionnaire.

Finally, the volume of query samples that were analyzed in this study was limited. In future work, after addressing the issues identified in the present evaluation through the expansion of the knowledge base and refinement of prompts, we plan to include more than 500 additional dialogs and re-evaluate the updated model. To enhance reproducibility and external validity, it will be necessary to pre-register multicenter, blinded studies that incorporate double coding with the calculation of inter-rater agreement, and ultimately conduct randomized controlled trials.

5 Conclusion

To the best of our knowledge, this is the first study worldwide to develop a specialized generative AI system to support the implementation of FBT and to demonstrate a very high rate of clinically appropriate responses in internal expert evaluations that simulate families implementing FBT. In particular, by incorporating domain-specific RAG and safety gates, we successfully established a model that achieved both alignment with the FBT principles and a high level of safety. These findings indicate the potential utility of the system as a continuous on-demand support tool for home use.

Future work should include expanding the FBT-concordant knowledge base, refining the escalation criteria, and further optimizing language generation. Therefore, it is necessary to prospectively evaluate the effectiveness and safety of the system through usability studies with caregivers who have experienced FBT, followed by randomized controlled trials.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

MnH: Resources, Writing – original draft, Formal analysis, Funding acquisition, Writing – review & editing, Project administration, Visualization, Methodology, Conceptualization, Data curation, Validation. JH: Data curation, Formal analysis, Writing – original draft, Supervision, Methodology, Validation, Software, Conceptualization, Funding acquisition, Writing – review & editing. AO: Supervision, Validation, Conceptualization, Writing – review & editing, Methodology. CT: Validation, Writing – review & editing. YS: Validation, Methodology, Writing – review & editing. CF: Writing – review & editing, Methodology. Mkh: Validation, Writing – review & editing. AS: Writing – review & editing, Validation. KT: Validation, Writing – review & editing, Software. RN: Validation, Software, Writing – review & editing. HK: Writing – review & editing, Conceptualization. YT: Validation, Writing – review & editing, Software. TO: Writing – review & editing, Supervision. HT: Writing – review & editing, Supervision.

Funding

The author(s) declared that financial support was received for this work and/or its publication. This research was funded by the Meiji Yasuda Mental Health Foundation and the Mishima Kaiun Memorial Foundation, Japan (both to M.H.), and by the Japan Science and Technology Agency [Grant Number: PSI2025_S2B71] and Hashimoto Foundation Inc., Japan (FY2024) (both to J.H.).

Conflict of interest

Author YT was employed by company NEC Corporation.

The remaining author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was used in the creation of this manuscript. *Napkin AI* was used to create Figure 2. Generative AI tools were used in the development of the system described in this manuscript. In addition, generative AI was used to create visual content and assist with English language editing of the manuscript text.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the

reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fdgth.2026.1759690/full#supplementary-material>

SUPPLEMENTARY FIGURE S1

References

- American Psychiatric Association. *Diagnostic & Statistical Manual of Mental Disorders, 5th ed., Text Revision (DSM-5-TR)*. Washington, DC: American Psychiatric Association Publishing (2022).
- Devoe DJ, Han A, Anderson A, Katzman DK, Patten SB, Soumbasis A, et al. The impact of the COVID-19 pandemic on eating disorders: a systematic review. *Int J Eat Disord.* (2023) 56(1):5–25. doi: 10.1002/eat.23704
- Reas DL, Rø G, Rø Ø. Trends in the observed versus expected incidence of eating disorders before, during, and after the COVID-19 pandemic: a national patient registry study. *Int J Eat Disord.* (2025) 58(8):1469–76. doi: 10.1002/eat.24443
- Steinhausen H-C. The outcome of anorexia nervosa in the 20th century. *Am J Psychiatry.* (2002) 159(8):1284–93. doi: 10.1176/appi.ajp.159.8.1284
- Pompili M, Mancinelli I, Girardi P, Ruberto A, Tatarelli R. Suicide in anorexia nervosa: a meta-analysis. *Int J Eat Disord.* (2004) 36(1):99–103. doi: 10.1002/eat.20011
- Lock J, Le Grange D, Agras WS, Moye A, Bryson SW, Jo B. Randomized clinical trial comparing family-based treatment with adolescent-focused individual therapy for adolescents with anorexia nervosa. *Arch Gen Psychiatry.* (2010) 67(10):1025–32. doi: 10.1001/archgenpsychiatry.2010.128
- Oshukova S, Suokas J, Nordberg M, Ålgars M. Effects of family-based treatment on adolescent outpatients treated for anorexia nervosa in the eating disorder unit of Helsinki university hospital. *J Eat Disord.* (2023) 11(1):154. doi: 10.1186/s40337-023-00879-9
- Rienecke RD, Le Grange D. The five tenets of family-based treatment for adolescent eating disorders. *J Eat Disord.* (2022) 10(1):60. doi: 10.1186/s40337-022-00585-y
- Iguchi T, Miyawaki D, Harada T, Ogiwara K. Introduction of family-based treatment to Japan with adaptations to optimize the cultural acceptability and advance current traditional treatments of adolescent anorexia nervosa. *Int J Eat Disord.* (2021) 54(1):17–9. doi: 10.1002/eat.23405
- Austin A, Anderson AG, Lee J, Vander Steen H, Savard C, Bergmann C, et al. Efficacy of eating disorder focused family therapy for adolescents with anorexia nervosa: a systematic review and meta-analysis. *Int J Eat Disord.* (2025) 58(1):3–36. doi: 10.1002/eat.24252
- Eating disorders: recognition and treatment. NICE. Available online at: <https://www.nice.org.uk/guidance/ng69> (Accessed November 13, 2025).
- Couturier J, Isserlin L, Norris M, Spettigue W, Brouwers M, Kimber M, et al. Canadian Practice guidelines for the treatment of children and adolescents with eating disorders. *J Eat Disord.* (2020) 8(1):4. doi: 10.1186/s40337-020-0277-8
- Crone C, Fochtmann LJ, Attia E, Boland R, Escobar J, Fornari V, et al. The American psychiatric association practice guideline for the treatment of patients with eating disorders. *Am J Psychiatry.* (2023) 180(2):167–71. doi: 10.1176/appi.ajp.23180001
- Dejong H, Broadbent H, Schmidt U. A systematic review of dropout from treatment in outpatients with anorexia nervosa. *Int J Eat Disord.* (2012) 45(5):635–47. doi: 10.1002/eat.20956
- Lavender KR. Rebooting “failed” family-based treatment. *Front Psychiatry.* (2020) 11:68. doi: 10.3389/fpsy.2020.00068
- Steinberg D, Perry T, Freestone D, Bohon C, Baker JH, Parks E. Effectiveness of delivering evidence-based eating disorder treatment via telemedicine for children, adolescents, and youth. *Eat Disord.* (2023) 31(1):85–101. doi: 10.1080/10640266.2022.2076334
- Hambleton A, Le Grange D, Kim M, Miskovic-Wheatley J, Touyz S, Maguire S. Delivering evidence-based treatment via telehealth for anorexia nervosa in rural health settings: a multi-site feasibility implementation study. *J Eat Disord.* (2024) 12(1):207. doi: 10.1186/s40337-024-01175-w
- Inoue T, Kosaka H, Okazaki R, Iida N, Isobe M, Inada S, et al. A multicenter retrospective study on anorexia Nervosa and avoidant/restrictive food intake disorder in Japan: the impact of the COVID-19 pandemic. *Jpn J Eating Disord.* (2023) 3(1):3–12. doi: 10.50983/jjed.3.1.3
- Maddox TM, Embi P, Gerhart J, Goldsack J, Parikh RB, Sarich TC. Generative AI in medicine - evaluating progress and challenges. *N Engl J Med.* (2025) 392(24):2479–83. doi: 10.1056/NEJMsb2503956
- Ayre J, Cvejic E, McCaffery KJ. Use of ChatGPT to obtain health information in Australia, 2024: insights from a nationally representative survey. *Med J Aust.* (2025) 222(4):210–2. doi: 10.5694/mja.252598
- Neha F, Bhati D, Shukla DK. Retrieval-augmented generation (RAG) in healthcare: a comprehensive review. *AI (Basel).* (2025) 6(9):226. doi: 10.3390/ai6090226
- Le Grange JLAD. *Treatment Manual for Anorexia Nervosa: A Family-based approach*. 2nd edn. New York: The Guilford Press (2015).
- Ganci M. *Survive FBT Skills Manual for Parents Undertaking Family Based Treatment(FBT) for Child and Adolescent Anorexia Nervosa*. Australia: LMD publishing (2016).
- Sakuta R. Hajimete no FBT Jissen Gaido [Introduction to Family-Based Treatment Practice Guide]. Eating Disorder Information Portal Site (n.d.). Available online at: <https://www.edportal.jp/pro/material.html> (Accessed August 1, 2025).
- Japanese Society of Psychosomatic Pediatrics. *Shouni Shinshin Igakukai Gaidorainshu [the Clinical Guidelines on Psychosomatic Diseases in Childhood]*. 3rd ed. Tokyo, Japan: Nankodo (2025).
- Hasei J, Hanzawa M, Nagano A, Maeda N, Yoshida S, Endo M, et al. Empowering pediatric, adolescent, and young adult patients with cancer utilizing generative AI chatbots to reduce psychological burden and enhance treatment engagement: a pilot study. *Front Digit Health.* (2025) 7:1543543. doi: 10.3389/fdgth.2025.1543543
- Busch F, Hoffmann L, Rueger C, van Dijk EHC, Kader R, Ortiz-Prado E, et al. Current applications and challenges in large language models for patient care: a systematic review. *Com Med.* (2025) 5(1):26. doi: 10.1038/s43856-024-00717-2
- Desage C, Patterson T, Wu Y, Bechtel A, Bharat A, Vaclavik D, et al. Illustrating a framework for assessing generative artificial intelligence-based conversational agents for mental health. *Psychiatry Res.* (2025) 356:116900. doi: 10.1016/j.psychres.2025.116900
- Wei Q, Yao Z, Cui Y, Wei B, Jin Z, Xu X. Evaluation of ChatGPT-generated medical responses: a systematic review and meta-analysis. *J Biomed Inform.* (2024) 151:104620. doi: 10.1016/j.jbi.2024.104620
- Waage D, Bryde A, Bentz M, Micali N. Family life and caregiving during family-based treatment (FBT) for anorexia nervosa: parent, patient, and clinician perspectives. *J Eat Disord.* (2025) 13(1):219. doi: 10.1186/s40337-025-01367-y
- Byrne CE, Accurso EC, Arnow KD, Lock J, Le Grange D. An exploratory examination of patient and parental self-efficacy as predictors of weight gain in adolescents with anorexia nervosa: eXPLORATORY EXAMINATION. *Int J Eat Disord.* (2015) 48(7):883–8. doi: 10.1002/eat.22376
- Hughes EK, Poker S, Bortz A, Yeo M, Telfer M, Sawyer SM. Adolescent and parent experience of care at a family-based treatment service for eating disorders. *Front Psychiatry.* (2020) 11:310. doi: 10.3389/fpsy.2020.00310
- Ohara C, Komaki G, Yamagata Z, Hotta M, Kamo T, Ando T. Factors associated with caregiving burden and mental health conditions in caregivers of patients with anorexia nervosa in Japan. *Biopsychosoc Med.* (2016) 10(1):21. doi: 10.1186/s13030-016-0073-5
- Schwarte R, Timmesfeld N, Dempfle A, Krei M, Egberts K, Jaite C, et al. Expressed emotions and depressive symptoms in caregivers of adolescents with

- first-onset anorexia nervosa-A long-term investigation over 2.5 years. *Eur Eat Disord Rev.* (2017) 25(1):44–51. doi: 10.1002/erv.2490
35. Ohara C, Nishizono-Maher A, Sekiguchi A, Sugawara A, Morino Y, Kawakami J, et al. Individualized peer support needs assessment for families with eating disorders. *Biopsychosoc Med.* (2023) 17(1):11. doi: 10.1186/s13030-023-00267-4
36. Whitney J, Murray J, Gavan K, Todd G, Whitaker W, Treasure J. Experience of caring for someone with anorexia nervosa: qualitative study. *Br J Psychiatry.* (2005) 187:444–9. doi: 10.1192/bjp.187.5.444
37. Thirunavukarasu AJ, Ting DSJ, Elangovan K, Gutierrez L, Tan TF, Ting DSW. Large language models in medicine. *Nat Med.* (2023) 29(8):1930–40. doi: 10.1038/s41591-023-02448-8
38. Dalle Grave R, Eckhardt S, Calugi S, Le Grange D. A conceptual comparison of family-based treatment and enhanced cognitive behavior therapy in the treatment of adolescents with eating disorders. *J Eat Disord.* (2019) 7(1):42. doi: 10.1186/s40337-019-0275-x
39. Saeed SA, Masters RM. Disparities in health care and the digital divide. *Curr Psychiatry Rep.* (2021) 23(9):61. doi: 10.1007/s11920-021-01274-4