

Using Baidu Search Queries and Large Language Models to Develop Person-Centred Guideline Questions for Allergic Rhinitis

Xin Luo, China

Fellowship type	Research Short-term Fellowship (3 months)
Fellowship period	1 June 2026 – 30 August 2026
Location	Porto, Portugal
Host institution	Faculty of Medicine, University of Porto (FMUP), Portugal
Host supervisor	Professor Bernardo Sousa-Pinto

1. Executive Summary

This EAACI Research Short-term Fellowship focused on developing and applying an artificial-intelligence-supported approach to identify person-centred clinical guideline questions for allergic rhinitis (AR) in mainland China. The project combined Baidu Index related-search data with large language model (LLM)-assisted text analysis and expert review, with the aim of capturing public information needs that may not be fully represented by conventional expert-driven guideline development.

Using six predefined AR-related keywords, 8,739 related-search records were retrieved for May 2025 to May 2026. After deduplication, 4,053 unique queries remained; 449 were considered relevant to AR management. Among these, 309 (68.8%) contained at least one identifiable PICO element. The LLM platform identified 163 potential foreground clinical questions; expert review retained 18 recommendation-ready queries, which were merged into 11 standardized guideline questions. Four corresponded to prioritized ARIA 2024–2025 questions, one to a non-prioritized ARIA question, and six were novel, mainly concerning traditional Chinese medicine and complementary approaches.

The fellowship therefore produced both a concrete research output—a full manuscript—and a reusable methodological framework for integrating digital information-seeking behaviour into guideline-question generation and regional guideline adaptation.

2. Fellowship Information

Project title	Using Baidu Search Queries and Large Language Models to Develop Person-Centred Guideline Questions for Allergic Rhinitis
Fellow	Xin Luo
Country	China
Category	EAACI Research Short-term Fellowship (3 months)

Dates	1 June 2026 – 30 August 2026
Location	Porto, Portugal
Host institution	Faculty of Medicine, University of Porto (FMUP), Portugal
Host supervisor	Professor Bernardo Sousa-Pinto

3. Research Questions Addressed and Rationale

Clinical practice guidelines depend on well-formulated, actionable questions, but conventional approaches may not fully capture the concerns of the broader patient population. Internet search behaviour provides a complementary window into real-world information needs. Previous work had shown that Google Trends data could contribute to ARIA guideline-question development, but that work mainly reflected Western search environments. The fellowship therefore addressed whether a comparable approach could be adapted to mainland China, where search behaviour and treatment interests may differ.

The project addressed four linked questions:

- Can Baidu Index search queries be used to identify person-centred, clinically actionable questions about AR management in mainland China?
- Can an LLM efficiently extract PICO elements and identify foreground clinical questions from large volumes of real-world search queries?
- How do questions derived from Chinese search behaviour compare with questions generated and prioritized in the ARIA 2024–2025 guideline process?
- Can this workflow provide a scalable method to support regional adaptation of international clinical practice guidelines?

4. Nature of the Research and Fellowship Activities

This was a methodological infodemiology and guideline-development study using publicly available, anonymised, aggregated search data. The fellowship integrated digital epidemiology, natural-language processing, clinical guideline methodology and expert review.

4.1 Data retrieval and preprocessing

Six representative AR-related keywords were used: “Rhinitis”, “Nasal congestion”, “Pollen”, “Anti-allergy medication”, “Allergic rhinitis”, and “Dust mite”. Related searches were retrieved weekly from Baidu Index for May 2025 to May 2026. Records were deduplicated and normalized before downstream analysis.

4.2 LLM-assisted analysis

An LLM-based platform was used to exclude queries irrelevant to AR management, extract Population, Intervention, Comparator and Outcome (PICO) elements, identify foreground clinical questions, and reformulate suitable searches into recommendation-oriented guideline questions.

4.3 Expert review and ARIA comparison

Candidate questions underwent manual review to ensure that they addressed AR management, involved actionable interventions, could be translated into recommendation-ready PICO questions, and were sufficiently specific for evidence synthesis. The final standardized questions were compared with the ARIA 2024–2025 question set and classified as prioritized, non-prioritized, or novel.

4.4 Research activities undertaken during the fellowship

The fellowship activities included refinement of the study design, adaptation of the previously established AI-supported workflow to Chinese-language Baidu data, data processing and interpretation, methodological and statistical discussions with the host team, expert review of candidate guideline questions, comparison with the ARIA 2024–2025 guideline-development process, and preparation and revision of the research manuscript. The supplied manuscript records Xin Luo and Professor Bernardo Sousa-Pinto as co-first authors and identifies Xin Luo as contributing to conceptualization/study design, methodology/statistical analysis, and original manuscript drafting.

5. Results

The project demonstrated the feasibility of combining population-level search data with LLM-assisted processing to generate candidate person-centred guideline questions for AR.

Study stage	Result
Related-search records retrieved	8,739
Unique queries after deduplication	4,053
Queries relevant to AR management	449
Queries with ≥ 1 identifiable PICO element	309 (68.8%)
Foreground clinical questions identified by LLM	163
Recommendation-ready queries retained after expert review	18
Final standardized guideline questions	11
Relationship to ARIA 2024–2025	4 prioritized; 1 non-prioritized; 6 novel

5.1 Main findings

Intervention was the most frequently identifiable PICO component. Most intervention-related searches were broad or unspecified, while specific interventions included traditional Chinese medicine, allergen immunotherapy, oral antihistamines, surgery, topical ointments, intranasal corticosteroids and laser therapy. Outcomes were dominated by symptom-oriented concerns, especially nasal congestion and runny nose.

Of the 11 final guideline questions, six (54.5%) were not represented in the ARIA 2024–2025 guideline-development process. These novel questions predominantly concerned traditional Chinese medicine and complementary approaches, including topical nasal ointments, ginger treatment, Cang'erzi-based nasal drops, herbal tea and laser therapy. This difference suggests that digital search data can reveal region-specific information-seeking patterns that may be missed when guideline questions are derived only from conventional expert processes.

6. Impact on Future Research

The findings support the use of digital search behaviour as a complementary source for identifying patient-centred questions before formal guideline prioritization. Importantly, the approach is not intended to replace expert panels; rather, it can broaden the evidence base used to define which questions should be considered.

For future research, the framework can be extended to other digital platforms—including social media, online health communities and news media—and applied across countries to compare regional information needs. It may also support GRADE-ADOLOPMENT by identifying locally relevant concerns for which adoption, adaptation or de novo recommendations may be required. Repeated analyses over time could further help determine whether public information needs change as treatments, disease awareness and healthcare policies evolve.

The fellowship also established a basis for continued collaboration between the University of Porto and the Chinese research team in AI-supported guideline methodology, digital epidemiology and person-centred evidence generation.

7. Outputs Generated from the Fellowship

A full research manuscript entitled “Using Baidu Search Queries and Large Language Models to Develop Person-Centred Guideline Questions for Allergic Rhinitis” was prepared from the fellowship project. The manuscript includes the complete study workflow, results, tables and interpretation, and acknowledges the generous support of the EAACI Research Fellowship. In the supplied version, no final journal publication citation is yet provided; therefore, the output is reported here as a manuscript generated from the fellowship rather than as a published article.

Future publications resulting from this exchange project will acknowledge the EAACI Fellowship in accordance with EAACI requirements.

8. Adaptations, Limitations and Lessons for Reproducibility

The study adapted an existing Google Trends-based workflow to the Chinese Baidu Index environment. This required adjustment of search terms because several terms used in the previous workflow did not have sufficient Baidu search volume; “Anti-allergy medication” was used instead, and “Dust mite” was added because of its relevance in China.

Several limitations should be considered when interpreting the results. Baidu Index permits only a one-year extraction window; search behaviour is not equivalent to healthcare need; aggregated nationwide data do not permit provincial analyses; and the LLM-assisted outputs, although followed by expert review, may miss or differently structure some relevant questions. These limitations also define clear priorities for future validation and replication.

9. Personal Reflection and Conclusion

The fellowship provided an opportunity to work at the intersection of allergy, guideline methodology, digital epidemiology and artificial intelligence. A key lesson was that real-world digital behaviour can add a distinct patient-facing perspective to guideline development, but meaningful use of these data requires rigorous preprocessing, transparent LLM-assisted methods and expert clinical review.

The experience strengthened my ability to translate large-scale, unstructured public data into clinically interpretable questions and to connect regional patient information needs with an international guideline framework. Going forward, I aim to improve this approach through broader data sources, stronger validation of LLM outputs, cross-country comparisons and closer integration with formal evidence synthesis and guideline adaptation.

Overall, the fellowship achieved a concrete research output while also creating a platform for continued international collaboration and future work on person-centred, locally responsive allergy guidelines.

10. Acknowledgements

I sincerely thank Professor Bernardo Sousa-Pinto for his supervision, scientific guidance and generous support throughout the fellowship, and the colleagues at the Faculty of Medicine, University of Porto, particularly the MEDCIDS/RISE@MEDCIDS team, for their methodological input and collaboration. I also thank Rafael José Vieira, Daniel Martinho-Dias and all collaborators who contributed to the project, as well as Professor Jean Bousquet and Professor Qintai Yang for their scientific guidance. I am especially grateful to EAACI for providing the Research Fellowship Award and making this international research exchange possible.

Xin Luo, M.D., Ph.D.

1. Department of Otolaryngology-Head and Neck Surgery, The Third Affiliated Hospital of Sun Yat-Sen University, Guangzhou 510630, China

2. Department of Allergy, The Third Affiliated Hospital of Sun Yat-Sen University, Guangzhou 510630, China

E-mail: luoxingz@126.com

Mobile/Whatsapp/Wechat: +86 13725406635