Scientific Research Publishing

Health, 2025, 17(11), 1427-1441
https://www.scirp.org/journal/health
ISSN Online: 1949-5005
ISSN Print: 1949-4998

Development and Validation of an Interactive, Bilingual Patient Decision Aid for Artificial Liver Support System Treatment in Acute-on-Chronic Liver Failure

Juan Wang, Meiling Zhang, Shan Ouyang, Miaoxia Chen, Lili Li*

Department of Infectious Diseases, The Third Affiliated Hospital of Sun Yat-sen University, Guangzhou, China
Email: *lilili@mail.sysu.edu.cn

How to cite this paper: Wang, J., Zhang, M.L., Ouyang, S., Chen, M.X. and Li, L.L. (2025) Development and Validation of an Interactive, Bilingual Patient Decision Aid for Artificial Liver Support System Treatment in Acute-on-Chronic Liver Failure. *Health*, **17**, 1427-1441.
https://doi.org/10.4236/health.2025.1711095

Received: October 30, 2025
Accepted: November 21, 2025
Published: November 24, 2025

Abstract

Background: Acute-on-chronic liver failure (ACLF) is a syndrome characterized by high short-term mortality. The decision to initiate an artificial liver support system (ALSS) is complex and preference-sensitive. Standardized tools to facilitate shared decision-making (SDM) in this critical context are lacking. **Objective:** This study aimed to systematically develop and validate an interactive, bilingual (Chinese/English) patient decision aid (PDA) for patients with ACLF considering ALSS, using a novel validation methodology employing large language models (LLMs). **Methods:** A three-phase, mixed-methods design was employed. Phase 1 (Scoping) involved a systematic literature review to identify core content. Phase 2 (Development) focused on content drafting and technical implementation of an interactive web-based prototype. Phase 3 (Validation) involved a two-pronged approach: 1) An innovative content validation process was conducted by systematically querying five distinct LLMs (GPT-4, Claude 3, Llama 3, Gemini Pro, and a domain-specific medical model) to simulate a multi-disciplinary expert review, assessing the PDA's content for accuracy, comprehensiveness, clarity, and neutrality. 2) Usability and acceptability were evaluated by 10 representative users (patients and family members) through a think-aloud protocol, the System Usability Scale (SUS), and semi-structured interviews. **Results:** A web-based PDA titled "Artificial Liver Decision Aid for ACLF" was successfully developed. The LLM-driven validation process resulted in a high degree of consensus on core medical facts and alignment with major clinical guidelines. The iterative querying process generated 17 actionable refinements, primarily enhancing the clarity of technical descriptions and adding nuance to risk-benefit statements. In the user testing, the mean SUS score was 87.5 (SD 6.2; range 75 - 95), cor-

responding to an "excellent" usability rating. Qualitative user feedback was overwhelmingly positive, highlighting the tool's clarity, ease of use, and the value of its interactive and bilingual features. **Conclusions:** We have developed and validated a high-quality PDA for the ALSS decision in ACLF, pioneering a novel and efficient LLM-based method for content validation. The tool demonstrates excellent usability and is a promising resource to support SDM. This study also presents a viable new paradigm for the rapid development of evidence-based patient-facing medical tools.

Keywords

Acute-on-Chronic Liver Failure, Artificial Liver Support System, Patient Decision Aid, Shared Decision Making, Tool Development, Large Language Models, Content Validation, Usability

1. Introduction

Acute-on-chronic liver failure (ACLF) is a devastating syndrome defined by the acute decompensation of chronic liver disease, accompanied by organ failure(s) and associated with high short-term mortality rates, often ranging from 50% to 90% [1]. Artificial liver support systems (ALSS) represent a critical therapeutic option, functioning as a bridge to liver transplantation or, in some cases, a bridge to recovery by temporarily performing the liver's detoxification and metabolic functions [2].

The decision of whether to initiate ALSS is a quintessential "preference-sensitive" decision, characterized by significant uncertainty and trade-offs [3]. While ALSS may offer a survival advantage, it is an invasive, costly, and resource-intensive therapy with no guaranteed benefit; some patients may not respond and still succumb to the disease [4]. This complex balance of potential benefits against substantial risks, costs, and procedural burdens frequently leads to profound decisional conflict for patients and their families [5].

Shared decision-making (SDM), a collaborative process where clinicians and patients weigh evidence and patient values to make a healthcare choice, is the recommended model for such situations [6]. Patient decision aids (PDAs) are evidence-based tools designed to facilitate SDM, shown to improve knowledge and reduce decisional conflict [7].

Despite the clear need, a specific, methodologically robust PDA for the ALSS decision in ACLF is absent. Traditional PDA development is resource-intensive, particularly the expert validation phase, which can be slow and challenging to coordinate [8]. Recent advancements in large language models (LLMs) present a novel opportunity to streamline and enhance this process [9] [10]. Therefore, the dual objectives of this study were: 1) to systematically develop an interactive, bilingual PDA for the ALSS decision, and 2) to pioneer and evaluate a novel validation methodology using a panel of LLMs, adhering to the principles of the Inter-

national Patient Decision Aid Standards (IPDAS).

2. Methods

2.1 Study Design and Framework

A multi-phase, mixed-methods study was conducted, guided by the IPDAS checklist. Key IPDAS criteria were systematically addressed, including a process for ensuring a balanced presentation of options through multi-model LLM review and a final expert check, and disclosing uncertainties in risk-benefit language, as detailed in Section 2.4.1. The study comprised three phases: 1) Scoping and content identification, 2) Prototype development and technical implementation, and 3) A hybrid validation process involving LLM-driven content validation and human-centered usability testing.

2.2. Phase 1: Scoping and Content Identification

A systematic literature search of PubMed, Embase, the Cochrane Library, and CNKI was conducted using terms including "acute-on-chronic liver failure", "artificial liver", "plasma exchange", "patient experience", and "decision making". We extracted data on ACLF definitions, ALSS modalities, efficacy, risks, and costs. This evidence formed the foundational knowledge base for the PDA content.

2.3. Phase 2: Prototype Development and Technical Implementation

Based on the evidence from Phase 1, the PDA content was drafted in plain language (6th - 8th grade reading level). A responsive web-based application was developed using HTML5, CSS3, and JavaScript, featuring a modular interface, bilingual (Chinese/English) functionality, an interactive decision balance sheet, and a values clarification exercise. The final source code is provided in **Appendix A**.

2.4. Phase 3: Validation

2.4.1 Content Validation via Large Language Model (LLM) Consultation

In a departure from traditional expert panels, we designed and executed a structured, two-round validation process simulating a multi-disciplinary expert review using five distinct LLMs: OpenAI's GPT-4, Anthropic's Claude 3 Opus, Meta's Llama 3, Google's Gemini Pro, and a specialized medical LLM (e.g., Med-PaLM 2, specified for context). This novel approach was designed to rapidly assess and refine the PDA's content for accuracy, comprehensiveness, clarity, and neutrality.

Process:

1. **Prompt Engineering:** A structured prompt protocol was developed (**Appendix B**). Each LLM was assigned a specific clinical persona (e.g., "You are a senior hepatologist specializing in ACLF", "You are a critical care nurse with 15 years of ICU experience", "You are a health literacy expert evaluating patient-facing materials").
2. **Round 1 (Independent Review):** The entire content of the PDA was fed to

each LLM with the structured prompt. The models were instructed to perform a line-by-line critique, identify inaccuracies, suggest clarifications, assess balance, and flag any missing crucial information.

3. **Synthesis:** All outputs from Round 1 were collated and analyzed. A human researcher synthesized the suggestions, identifying points of consensus, divergence, and unique insights from each model.
4. **Round 2 (Consensus Building):** A summary of the synthesized feedback and points of divergence was fed back to the LLMs. They were prompted to review the critiques from the other (anonymized) "experts" and either reaffirm their position, modify it, or reach a consensus, simulating a Delphi-like process.

Human Oversight: To mitigate the risk of LLM errors, a final human oversight step was crucial. Following the synthesis of LLM suggestions, the corresponding author, a clinical expert in hepatology, reviewed all high-consensus and divergent suggestions before their incorporation. The expert then performed a final line-by-line review of the complete, revised PDA content to ensure all changes were medically sound, appropriate, and contextually correct before proceeding to user testing.

2.4.2. Usability and Acceptability Testing

To assess user-friendliness, we recruited a convenience sample of 10 end-users (5 patients with chronic liver disease, 5 family members). Recruitment occurred at the outpatient clinic of The Third Affiliated Hospital of Sun Yat-sen University. Eligibility criteria included: adults (≥18 years) who were either diagnosed with chronic liver disease (but not in an acute ACLF state, to avoid undue distress) or were a family member of such a patient, and were able to provide informed consent and communicate in Chinese. This convenience sampling approach was deemed appropriate for this stage of usability testing, which focuses on interface clarity and user experience rather than clinical outcomes. Each participant engaged in a "think-aloud" session while using the PDA. Subsequently, they completed the 10-item System Usability Scale (SUS) [11] (**Appendix C**) and a brief semi-structured interview to gather qualitative feedback.

2.5. Ethical Considerations

The study protocol was approved by the Institutional Review Board of The Third Affiliated Hospital of Sun Yat-sen University (RG2023-170-03). All human participants (users) provided written informed consent.

3. Results

3.1 Final PDA Prototype

The final product is a stand-alone, interactive, bilingual web-based application. It comprises five sections: 1) Introduction to ACLF; 2) Treatment Options; 3) Interactive Pros and Cons; 4) Values Clarification; and 5) A personalized summary page.

3.2 Content Validation by LLMs

The two-round LLM validation process was completed within 48 hours. A high degree of consensus (>4 of 5 models) was observed for core medical definitions, treatment mechanisms, and major risk categories. A total of 17 distinct, actionable suggestions for content refinement were generated and subsequently incorporated. Thematic analysis of the LLM feedback revealed three key areas of improvement:

1. **Enhanced Precision:** LLMs suggested replacing general terms with more specific information. For example, "lying in bed for a long time" was refined to "lying relatively still for 2 - 4 hours per session".
2. **Nuanced Risk-Benefit Language:** The models consistently recommended softening deterministic language. For instance, "ALSS increases survival rate" was revised to "ALSS may increase survival rate for certain patients", better reflecting clinical evidence uncertainty.
3. **Addition of Practical Information:** Suggestions included adding a brief note about the need for family support during the lengthy treatment sessions and clarifying the difference between ALSS and kidney dialysis, both of which were integrated.

Points of divergence among models, such as the exact statistical range for survival benefit, were used to guide the final text towards more cautious and generalized phrasing, explicitly stating that outcomes vary significantly. To better illustrate the quality of the feedback, more specific examples of actionable suggestions included:

1. **Refining Vague Descriptions:** Changing "The treatment takes a long time" to "Each ALSS session typically requires you to be relatively still for 2 - 4 hours".
2. **Quantifying where appropriate:** Modifying "ALSS has a high cost" to include a note: "Costs can be substantial and vary by region and the specific ALSS mode used; please discuss the estimated financial impact with the hospital's billing department and your care team."
3. **Adding Patient-Centric Clarifications:** One model, role-playing as a health literacy expert, suggested adding the explicit comparison: "You can think of ALSS as a form of 'liver dialysis,' similar in concept to the kidney dialysis many people are familiar with, but for the liver."

3.3. User Characteristics

The 10 representative users included 5 patients (3 male, 2 female) and 5 family members (2 male, 3 female). The demographic and clinical characteristics of the sample are summarized in Table 1.

Table 1. Demographic and clinical characteristics of usability testing participants (N = 10).

Characteristic	Category	n (%)
User Type	Patient with Chronic Liver Disease	5 (50%)
	Family Member	5 (50%)

Continued

Age (years), range		35-65
Gender	Male	5 (50)
	Female	5 (50)
Education Level	Middle school or below	3 (30)
	High school/Technical school	4 (40)
	College degree or above	3 (30)
Diagnosis (Patients only, n = 5)	Hepatitis B-related Cirrhosis	3 (60)
	Alcoholic Cirrhosis	1 (20)
	Other	1 (20)

3.4. Usability and Acceptability

All 10 users completed the usability testing. The mean SUS score was 87.5 (SD 6.2; range 75 - 95), indicating "excellent" usability and placing it in the 90 - 95th percentile [11]. Qualitative feedback was highly positive, with themes of clarity, the value of interactivity, and empowerment through family communication via the bilingual feature emerging consistently.

4. Discussion

This study reports on the successful development of a high-quality PDA for the ALSS decision in ACLF. More significantly, it pioneers a novel, efficient, and robust methodology for content validation using a panel of LLMs. This approach addresses a major bottleneck in traditional PDA development, offering a scalable and rapid alternative to the initial stages of expert review.

The principal strength of the LLM validation method is its speed and breadth. It can synthesize vast amounts of published literature and guidelines in seconds, providing a comprehensive check for accuracy and completeness that may surpass that of a single human expert. The use of multiple, diverse models and persona-based prompting simulates a multi-disciplinary panel, allowing for cross-validation and reducing the risk of individual model bias. Our results show this method yielded specific, high-quality refinements comparable to those expected from a human panel.

The methodological rigor of our study, combining this innovative LLM validation with gold-standard human-centered usability testing, ensures the final product is both evidence-aligned and user-friendly. The high SUS score and positive qualitative feedback confirm the PDA's potential to be a valuable clinical tool.

We acknowledge several limitations. First, the primary limitation is the reliance on LLMs for content validation. LLMs can "hallucinate" or provide outdated information and lack real-world clinical experience and nuance [12]. To mitigate this, we used multiple models, cross-verified outputs, and had a clinical expert (the corresponding author) perform a final review of all AI-suggested changes, as detailed in our methods. We position this method as a powerful accelerator for

initial development and refinement, not a complete replacement for final human oversight. Second, the user testing was conducted with a small, single-center sample, which may limit generalizability. Third, this study does not assess the PDA's clinical efficacy.

The clear next step is a randomized controlled trial (RCT) to evaluate the PDA's effectiveness in a clinical setting. Specifically, this future RCT should measure primary outcomes such as a change in the Decisional Conflict Scale (DCS) score, and secondary outcomes including patient knowledge of ALSS, alignment between patient values and the chosen treatment (decision-value congruence), and patient satisfaction with the decision-making process.

5. Conclusion

We have rigorously developed and validated an interactive, bilingual PDA for the ALSS decision in ACLF using a novel combination of LLM-driven content refinement and end-user testing. The tool is content-sound, highly usable, and ready for clinical evaluation. This study demonstrates that a structured, multi-model LLM consultation can serve as a powerful and efficient alternative to traditional methods for the initial content validation of patient-facing medical tools, heralding a new paradigm in the agile development of digital health interventions.

Funding

This study was supported by the Nursing Research Fund of the Third Affiliated Hospital of Sun Yat-sen University (Project No. 2022HLZD02).

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Choudhury, A., Kulkarni, A.V., Arora, V., Soin, A.S., Dokmeci, A.K., Chowdhury, A., *et al.* (2025) Acute-On-Chronic Liver Failure (ACLF): The 'Kyoto Consensus'—Steps from Asia. *Hepatology International*, **19**, 1-69. https://doi.org/10.1007/s12072-024-10773-4

[2] Allhoff, B., Neumann-Haefelin, C. and Kasper, P. (2025) Acute-on-Chronic Liver Failure—Current Management and Future Perspectives. *Biomedicines*, **13**, Article 2193. https://doi.org/10.3390/biomedicines13092193

[3] Ocskay, K., Kanjo, A., Gede, N., Szakács, Z., Pár, G., Erőss, B., *et al.* (2021) Uncertainty in the Impact of Liver Support Systems in Acute-On-Chronic Liver Failure: A Systematic Review and Network Meta-Analysis. *Annals of Intensive Care*, **11**, Article No. 10. https://doi.org/10.1186/s13613-020-00795-0

[4] Saliba, F., Bañares, R., Larsen, F.S., Wilmer, A., Parés, A., Mitzner, S., *et al.* (2022) Artificial Liver Support in Patients with Liver Failure: A Modified DELPHI Consensus of International Experts. *Intensive Care Medicine*, **48**, 1352-1367. https://doi.org/10.1007/s00134-022-06802-1

[5] Scherer, K.A., Büdenbender, B., Blum, A.K., Grüne, B., Kriegmair, M.C., Michel, M.S.,

et al. (2025) Power Asymmetry and Embarrassment in Shared Decision-Making: Predicting Participation Preference and Decisional Conflict. *BMC Medical Informatics and Decision Making*, **25**, Article No. 120. https://doi.org/10.1186/s12911-025-02938-4

[6] Elwyn, G., Frosch, D., Thomson, R., Joseph-Williams, N., Lloyd, A., Kinnersley, P., *et al.* (2012) Shared Decision Making: A Model for Clinical Practice. *Journal of General Internal Medicine*, **27**, 1361-1367. https://doi.org/10.1007/s11606-012-2077-6

[7] Stacey, D., Légaré, F., Lewis, K., Barry, M.J., Bennett, C.L., Eden, K.B., *et al.* (2017) Decision Aids for People Facing Health Treatment or Screening Decisions. *Cochrane Database of Systematic Reviews*, No. 4, cd001431. https://doi.org/10.1002/14651858.cd001431.pub5

[8] Elwyn, G., O'Connor, A., Stacey, D., *et al.* (2006) Developing a Quality Criteria Framework for Patient Decision Aids: Online International Delphi Consensus Process. *BMJ: British Medical Journal*, **333**, 417.

[9] Hinds, J. and Joinson, A.N. (2024) Digital Data and Personality: A Systematic Review and Meta-Analysis of Human Perception and Computer Prediction. *Psychological Bulletin*, **150**, 727-766. https://doi.org/10.1037/bul0000430

[10] Wiest, I.C., Bhat, M., Clusmann, J., Schneider, C.V., Jiang, X. and Kather, J.N. (2025) Large Language Models for Clinical Decision Support in Gastroenterology and Hepatology. *Nature Reviews Gastroenterology & Hepatology*, **22**, 773-787. https://doi.org/10.1038/s41575-025-01108-1

[11] Bangor, A., Kortum, P. and Miller, J. (2009) Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale. *Journal of Usability Studies*, **4**, 114-123.

[12] Hobensack, M., von Gerich, H., Vyas, P., Withall, J., Peltonen, L., Block, L.J., *et al.* (2024) A Rapid Review on Current and Potential Uses of Large Language Models in Nursing. *International Journal of Nursing Studies*, **154**, Article ID: 104753. https://doi.org/10.1016/j.ijnurstu.2024.104753

Appendix A: Final Source Code of the Patient Decision Aid

```
<!DOCTYPE html>
<html lang="zh">
<head>
    <meta charset="UTF-8">
    <meta name="viewport" content="width=device-width, initial-scale=1.0">
    <title>ACLF 人工肝治疗决策辅助工具 - Artificial Liver Decision Aid</title>
    <style>
        :root { --primary-color: #2980b9; --secondary-color: #ecf0f1; --text-color: #34495e; --card-bg: #ffffff; --border-color: #bdc3c7; } body { font-family: -apple-system, BlinkMacSystemFont, "Segoe UI", Roboto, "Helvetica Neue", Arial, "Noto Sans", sans-serif; margin: 0; padding: 20px; background-color: var(--secondary-color); color: var(--text-color); line-height: 1.6; } .container { max-width: 900px; margin: auto; background-color: var(--card-bg); padding: 20px 30px; border-radius: 8px; box-shadow: 0 4px 8px rgba(0,0,0,0.1); } header { text-align: center; border-bottom: 2px solid var(--primary-color); padding-bottom: 15px; margin-bottom: 25px; position: relative; } header h1 { color: var(--primary-color); margin: 0; } .lang-switch { position: absolute; top: 0; right: 0; } .lang-switch button { background-color: var(--primary-color); color: white; border: none; padding: 8px 12px; border-radius: 5px; cursor: pointer; font-size: 14px; transition: background-color 0.3s; } .lang-switch button:hover { opacity: 0.9; } .lang-switch button.active { background-color: #1a5276; font-weight: bold; } .card { background: var(--card-bg); border: 1px solid var(--border-color); border-radius: 8px; padding: 20px; margin-bottom: 20px; box-shadow: 0 2px 4px rgba(0,0,0,0.05); } h2 { color: var(--primary-color); border-bottom: 1px solid var(--secondary-color); padding-bottom: 10px; } .options { display: grid; grid-template-columns: 1fr 1fr; gap: 20px; } .option-card { border: 2px solid var(--border-color); border-radius: 8px; padding: 15px; text-align: center; } .option-card h3 { margin-top: 0; } .pros-cons-grid { display: grid; grid-template-columns: 1fr 1fr; gap: 15px; } .pros, .cons { padding: 15px; border-radius: 5px; } .pros { background-color: #e8f6f3; border-left: 4px solid #1abc9c; } .cons { background-color: #fdedec; border-left: 4px solid #e74c3c; } .pros h4, .cons h4 { margin-top: 0; } .pros h4 { color: #16a085; } .cons h4 { color: #c0392b; } ul { padding-left: 20px; } li { margin-bottom: 10px; } .values-clarification-grid { display: grid; grid-template-columns: repeat(auto-fit, minmax(250px, 1fr)); gap: 15px; } .value-item { background-color: #fdf2e9; padding: 15px; border-radius: 5px; border-left: 4px solid #e67e22; } .decision-balance-table { width: 100%; border-collapse: collapse; margin-top: 20px; } .decision-balance-table th, .decision-balance-table td { border: 1px solid var(--border-color); padding: 12px; text-align: left; } .decision-balance-table th { background-color: var(--secondary-color); color: var(--primary-color); } .important-btn { background-color: #f1c40f; color: var(--text-color); border: none; padding: 5px 10px; border-radius: 5px; cursor: pointer; font-
```

size: 12px; margin-left: 10px; float: right; } .important-btn.clicked { background-color: #f39c12; font-weight: bold; } #summary-card li { list-style-type: '✅ '; } @media (max-width: 768px) { .options, .pros-cons-grid, .values-clarification-grid { grid-template-columns: 1fr; } .lang-switch { position: static; text-align: center; margin-top: 10px; } }
</style>
</head>
<body>
<header><h1 data-lang-zh="ACLF 人工肝治疗决策辅助工具" data-lang-en="Artificial Liver Decision Aid for ACLF"></h1> <button id="btn-zh" onclick="switchLanguage('zh')"> 中 文 </button><button id="btn-en" onclick="switchLanguage('en')">English</button> </header> <h2 data-lang-zh="第一步：了解您的疾病" data-lang-en="Step 1: Understand Your Condition"></h2><p data-lang-zh="您被诊断为“慢加急性肝衰竭”(ACLF)。这意味着您原有的慢性肝病（如肝硬化）突然严重恶化，导致肝脏功能急剧下降，并可能影响其他器官（如肾脏、大脑）的功能。这是一个非常危重的状况，需要立即进行积极治疗。" data-lang-en="You have been diagnosed with 'Acute-on-Chronic Liver Failure' (ACLF). This means your existing chronic liver disease (like cirrhosis) has suddenly worsened, leading to a rapid decline in liver function and potentially affecting other organs (like kidneys, brain). This is a very critical condition that requires immediate and aggressive treatment."> <h2 data-lang-zh="第二步：了解您的治疗选择" data-lang-en="Step 2: Understand Your Treatment Options"></h2><p data-lang-zh="针对 ACLF，目前主要有两种治疗策略。您的医疗团队会与您讨论哪一种更适合您目前的情况。" data-lang-en="For ACLF, there are currently two main treatment strategies. Your medical team will discuss with you which one is more suitable for your current situation."> <h3 data-lang-zh="选择 A：标准内科治疗 (SMT)" data-lang-en="Option A: Standard Medical Therapy (SMT)"></h3><p data-lang-zh="这是基础治疗，包括使用药物（如白蛋白、抗生素）、营养支持、处理并发症（如腹水、肝性脑病）等，目标是稳定病情，支持身体自身修复。" data-lang-en="This is the foundational treatment, including medications (like albumin, antibiotics), nutritional support, and managing complications (like ascites, hepatic encephalopathy). The goal is to stabilize the condition and support the body's own repair mechanisms."> <h3 data-lang-zh="选择 B：人工肝支持系统 (ALSS) + SMT" data-lang-en="Option B: Artificial Liver Support System (ALSS) + SMT"></h3><p data-lang-zh="在标准内科治疗的基础上，增加人工肝治疗。这是一种类似“透析”的技术，通过一台机器帮助您的身体清除毒素，暂时替代部分肝脏功能，为肝脏恢复或等待肝移植创造机会。"data-lang-en="In addition to standard medical therapy, artificial liver treatment

is added. It's a technology similar to 'dialysis' that uses a machine to help your body clear toxins, temporarily replacing some liver functions, creating an opportunity for liver recovery or bridging to a liver transplant."></p> <h2 data-lang-zh="第三步：权衡利弊" data-lang-en="Step 3: Weigh the Pros and Cons"></h2><p data-lang-zh="每个选择都有其优点和缺点。请仔细阅读，并点击右侧的“对我很重要”按钮，标记出您最在意的项目。这些项目将会汇总到最后的决策平衡表中。"data-lang-en="Each option has its pros and cons. Please read them carefully and click the 'Important to me' button on the right to mark the items you care about most. These items will be summarized in the final decision balance table."></p><h3 data-lang-zh="选择 A：标准内科治疗 (SMT)" data-lang-en="Option A: Standard Medical Therapy (SMT)"></h3> <h4 data-lang-zh="优点" data-lang-en="Pros"></h4>
<li data-id="smt-pro-1" data-lang-zh="无创伤性：避免了与置管相关的风险（如出血、感染）。" data-lang-en="Non-invasive: Avoids risks associated with catheter insertion (e.g., bleeding, infection).">
<li data-id="smt-pro-2" data-lang-zh="费用相对较低：主要为药物和支持治疗的费用。" data-lang-en="Relatively lower cost: Mainly consists of medication and supportive care costs."> <h4 data-lang-zh="缺点" data-lang-en="Cons"></h4>
<li data-id="smt-con-1" data-lang-zh="效果可能有限：对于非常严重的 ACLF，可能无法有效清除体内大量毒素，逆转病情。" data-lang-en="May have limited effect: For very severe ACLF, it may not effectively clear the large amount of toxins in the body to reverse the condition.">
<li data-id="smt-con-2" data-lang-zh="死亡率可能更高：对于某些中重度患者，不进行 ALSS 的短期死亡率可能更高。" data-lang-en="Potentially higher mortality: For some moderate to severe patients, the short-term mortality rate may be higher without ALSS."> <h3 data-lang-zh="选择 B：人工肝支持系统(ALSS) + SMT" data-lang-en="Option B: Artificial Liver Support System (ALSS) + SMT" style="margin-top: 20px;"></h3> <h4 data-lang-zh="优点" data-lang-en="Pros"></h4>
<li data-id="alss-pro-1" data-lang-zh="可能提高生存率：研究表明，对于特定患者，ALSS 可能降低短期死亡率，为肝脏恢复或肝移植赢得时间。"data-lang-en="May increase survival rate: Studies show that for specific patients, ALSS may reduce short-term mortality, buying time for liver recovery or transplant.">
<li data-id="alss-pro-2" data-lang-zh="快速改善症状：能有效清除胆红素、氨等毒素，可能快速改善黄疸和肝性脑病等症状。"data-lang-en="Rapid symptom improvement: Can effectively clear toxins like bilirubin and ammonia, potentially improving symptoms like jaundice and hepatic encephalopathy quickly."> <h4 data-lang-zh="缺点"data-lang-en="Cons"></h4>
<li data-id="alss-con-1" data-lang-zh="侵入性操作：需要在颈部或大腿的大血管内置入一根导管，存在出血、感染、血栓形成的风险。"data-lang-en="Invasive procedure: Requires inserting a catheter into a large blood vessel in the neck or thigh, with risks of bleeding, infection, and thrombosis.">
<li data-id="alss-con-2" data-lang-zh="费用高昂：单次治疗费

用较高，且通常需要多次治疗，对家庭经济是巨大考验。" data-lang-en="High cost: The cost per session is high, and multiple sessions are usually required, posing a significant financial burden on the family.">
<li data-id="alss-con-3" data-lang-zh="疗效不确定：并非对所有患者都有效，部分患者即使接受了治疗，病情仍可能继续恶化。"data-lang-en="Uncertain efficacy: It is not effective for all patients, and some may continue to deteriorate despite treatment.">
<li data-id="alss-con-4" data-lang-zh="治疗过程可能不适：需要在病床上平躺数小时，可能出现血压波动、寒战等反应。"data-lang-en="Potential discomfort during treatment: Requires lying relatively still for 2-4 hours per session, and may cause blood pressure fluctuations, chills, etc."> <h2 data-lang-zh="第四步：思考什么对您最重要？"data-lang-en="Step 4: Think About What's Most Important to You"></h2> <strong data-lang-zh="关于家庭负担："data-lang-en="About Family Burden:"> <h2 data-lang-zh="第五步：您的决策平衡表" data-lang-en="Step 5: Your Decision Balance Sheet"></h2><p data-lang-zh="这是根据您刚才标记的“对我很重要”的项目，为您生成的个人决策总结。您可以把这个总结作为与医生和家人讨论的基础。"data-lang-en="This is a personal decision summary generated based on the items you marked as 'Important to me'. You can use this summary as a basis for discussion with your doctor and family."> <table class="decision-balance-table" id="decision-balance-table"></table><p id="no-selection-msg" data-lang-zh="您还没有选择任何重要的项目。请返回第三步，点击您关心的项目旁的“对我很重要”按钮。"data-lang-en="You haven't selected any important items yet. Please go back to Step 3 and click the 'Important to me' button next to the items you care about.">

<script>

const translations = {}; function collectTranslations() { document.querySelectorAll('[data-lang-zh]').forEach(el => { const key = el.tagName + '_' + (el.id || Array.from(el.attributes).map(a => a.name).join('_')); translations[key] = { zh: el.getAttribute('data-lang-zh'), en: el.getAttribute('data-lang-en') }; el.dataset.translationKey = key; }); } function switchLanguage(lang) { document.documentElement.lang = lang; docu-

```
ment.querySelectorAll('[data-translation-key]').forEach(el => { const key = el.dataset.translationKey; if (translations[key] && translations[key][lang]) { el.innerHTML = translations[key][lang]; } }); document.getElementById('btn-zh').classList.toggle('active', lang === 'zh'); document.getElementById('btn-en').classList.toggle('active', lang === 'en'); generateDecisionBalanceTable(lang); } let importantItems = new Set(); function toggleImportant(btn, id) { btn.classList.toggle('clicked'); if (importantItems.has(id)) { importantItems.delete(id); } else { importantItems.add(id); } const currentLang = document.documentElement.lang || 'zh'; generateDecisionBalanceTable(currentLang); } function generateDecisionBalanceTable(lang) { const table = document.getElementById('decision-balance-table'); const noSelectionMsg = document.getElementById('no-selection-msg'); let tableHTML = `<thead><tr><th>${lang === 'zh' ? '治疗选择' : 'Treatment Option'}</th><th>${lang === 'zh' ? '对我重要的优点' : 'Pros that Matter to Me'}</th><th>${lang === 'zh' ? '对我重要的缺点' : 'Cons that Matter to Me'}</th></tr></thead><tbody><tr><td><strong>${lang === 'zh' ? '标准内科治疗 (SMT)' : 'Standard Medical Therapy (SMT)'}</strong></td><td id="summary-smt-pros"></td><td id="summary-smt-cons"></td></tr><tr><td><strong>${lang === 'zh' ? '人工肝 (ALSS) + SMT' : 'Artificial Liver (ALSS) + SMT'}</strong></td><td id="summary-alss-pros"></td><td id="summary-alss-cons"></td></tr></tbody>`; table.innerHTML = tableHTML; let hasSelection = false; importantItems.forEach(id => { hasSelection = true; const element = document.querySelector(`[data-id="${id}"]`); if (element) { const text = element.getAttribute(`data-lang-${lang}`); const li = `<li>${text}</li>`; if (id.startsWith('smt-pro')) document.getElementById('summary-smt-pros').innerHTML += li; if (id.startsWith('smt-con')) document.getElementById('summary-smt-cons').innerHTML += li; if (id.startsWith('alss-pro')) document.getElementById('summary-alss-pros').innerHTML += li; if (id.startsWith('alss-con')) document.getElementById('summary-alss-cons').innerHTML += li; } }); if (hasSelection) { table.style.display = 'table'; noSelectionMsg.style.display = 'none'; } else { table.style.display = 'none'; noSelectionMsg.style.display = 'block'; } } function addImportantButtons() { document.querySelectorAll('li[data-id]').forEach(li => { const id = li.dataset.id; const btn = document.createElement('button'); btn.className = 'important-btn'; btn.setAttribute('data-lang-zh', '对我很重要'); btn.setAttribute('data-lang-en', 'Important to me'); btn.onclick = () => toggleImportant(btn, id); li.appendChild(btn); }); } document.addEventListener('DOMContentLoaded', () => { addImportantButtons(); collectTranslations(); switchLanguage('zh'); });
</script>
</body>
</html>
```

Appendix B: Structured Prompt Protocol for LLM Validation (Example)

Objective: To systematically evaluate the content of a patient decision aid (PDA) for ALSS in ACLF. The following prompt structure was used for each LLM.

--- START OF PROMPT ---

Persona Assignment:

You are a world-leading [Assigned Persona: e.g., 'Hepatologist specializing in liver failure', 'Critical Care Nurse with 15 years ICU experience', 'Health Literacy Expert focused on patient communication', 'Medical Ethicist specializing in end-of-life decisions', 'Biostatistician focused on clinical trial evidence']. You are part of a simulated expert panel reviewing a new patient decision aid.

Task:

Critically evaluate the following text from a patient decision aid for accuracy, comprehensiveness, clarity, and neutrality, from the perspective of your assigned persona. Provide your feedback in a structured format.

Evaluation Criteria:

1. **Accuracy:** Is the medical information factually correct according to the latest clinical evidence and guidelines (e.g., APASL, EASL)? Identify any statement that is inaccurate, misleading, or outdated.
2. **Comprehensiveness:** Is any critical information missing? Are there important risks, benefits, alternatives, or uncertainties that a patient and their family absolutely must know but are not included?
3. **Clarity & Plain Language:** Is the language simple, clear, and free of jargon? Is it likely to be understood by a patient with a middle-school education level? Suggest alternative wording for any complex sentences.
4. **Neutrality & Balance:** Is the information presented in a balanced, unbiased way? Does it avoid steering the patient toward one option over another? Flag any loaded or emotionally charged language.

Content for Review:

[Insert a section of the PDA content here, e.g., the "Pros and Cons of ALSS" section]

Output Format:

Please provide your feedback as a list of specific, actionable points. For each point, state:

- **Item:** The specific text you are critiquing.
- **Issue:** The problem you have identified (e.g., "Inaccurate," "Lacks clarity," "Biased language," "Missing information").
- **Rationale:** A brief explanation of why it is an issue from your persona's perspective.
- **Suggestion:** A concrete recommendation for revision (e.g., "Change 'X' to 'Y'," "Add a sentence explaining Z").

If you find no issues with a section, state "No issues found in this section."

--- END OF PROMPT ---

Appendix C: The System Usability Scale (SUS)

Instructions: For each of the following statements, please mark one box that best represents your reactions to the decision aid tool.

Statement	Strongly Disagree (1)	(2)	(3)	(4)	Strongly Agree (5)
1. I think that I would like to use this system frequently.	☐	☐	☐	☐	☐
2. I found the system unnecessarily complex.	☐	☐	☐	☐	☐
3. I thought the system was easy to use.	☐	☐	☐	☐	☐
4. I think that I would need the support of a technical person to be able to use this system.	☐	☐	☐	☐	☐
5. I found the various functions in this system were well integrated.	☐	☐	☐	☐	☐
6. I thought there was too much inconsistency in this system.	☐	☐	☐	☐	☐
7. I would imagine that most people would learn to use this system very quickly.	☐	☐	☐	☐	☐
8. I found the system very cumbersome to use.	☐	☐	☐	☐	☐
9. I felt very confident using the system.	☐	☐	☐	☐	☐
10. I needed to learn a lot of things before I could get going with this system.	☐	☐	☐	☐	☐

Scoring: For odd-numbered items, subtract 1 from the user's response. For even-numbered items, subtract the user's response from 5. Sum the scores for all 10 items and multiply by 2.5 to obtain the overall SUS score (0 - 100).