Progress in
Medical Education

Zentime

RESEARCH ARTICLE

The practice and exploration of argument-based pedagogy through academic controversy in an eight-year medical immunology program: A case study of the "Nobel Prize Controversy Involving Liping Chen"

Liyuan Zhao[1], Yijie Tao[2], Min Zhang[3], Sheng Xu[1]

[1]National Key Laboratory of Immunology & Inflammation, Naval Medical University, Shanghai 200433, China.
[2]Department of Physiology of Anesthesia, School of Anesthesiology, Naval Medical University, Shanghai 200433, China.
[3]Department of Laboratory Animal Sciences, School of Basic Medicine, Naval Medical University, Shanghai 200433, China.

Corresponding authors: Sheng Xu and Min Zhang.

Address correspondence to: Sheng Xu, National Key Laboratory of Immunology & Inflammation, Naval Medical University, 800 Xiangyin Road, Yangpu District, Shanghai 200433, China.
Tel: +86-13564932474.
E-mail: x.xusheng@163.com.
Min Zhang, Department of Laboratory Animal Sciences, School of Basic Medicine, Naval Medical University, 800 Xiangyin Road, Yangpu District, Shanghai 200433, China.
Tel: +86-18764201333.
E-mail: zhangmin_vet@126.com.

Received October 28, 2025; Accepted December 4, 2025; Published April 1, 2026

DOI: 10.61189/466869iuwenb

Abstract

Eight-year medical programs aim to train physician-scientists capable of critically evaluating evidence and navigating complex ethical dilemmas. Medical Immunology, straddling multiple disciplines, can nurture these abilities. We hypothesized that a structured academic controversy (SAC), centered on a high-profile scientific dispute, could simultaneously reinforce conceptual understanding and foster higher-order thinking. A two-hour, debate-based seminar was conducted one week following the lectures on B-lymphocyte and antibody-mediated immunity. The case focused on the "Lieping Chen Nobel Prize controversy" concerning priority in the discovery of the PD-1/PD-L1 pathway. A pre-class micro-package—including an original article from *Proceedings of the National Academy of Sciences of the United States of America*, excerpts from the Nobel white paper, and a three-minute animation—was provided to prime students. In class, students were randomly assigned to pro or con teams and engaged in a 50-minute timed debate, followed by rebuttals after switching sides. Real-time scoring rubrics, a 6-item Likert scale (assessing critical thinking and ethical sensitivity), and 60-second post-class audio reflections provided multi-source evaluation data. The mean critical-thinking score rose from 3.2 to 4.1 ($p<0.01$), and the ethical-sensitivity score from 3.4 to 3.8 ($p<0.05$). A one-month transfer test showed that 83% of students accurately applied PD-1/PD-L1 concepts to novel immunological contexts, while 47% extended their critical inquiry to new targets (e.g., CD47, LAG-3). Qualitative analysis revealed an increased appreciation for collaborative credit and a decrease in ad hominem language. In conclusion, a single, tightly integrated 2-hour SAC debate significantly enhanced conceptual mastery, critical appraisal, and ethical reasoning without requiring additional curriculum time. This model can be scaled to other contentious scientific discoveries as a practical method for developing evidence-based and ethically-minded physician-scientists.

Keywords: Eight-year medical program, Medical Immunology, Debate-based instruction, PD-1/PD-L1, Lieping Chen, Nobel Prize, Critical thinking

Highlights

- A single two-hour, debate-based seminar embedded within an eight-year medical immunology curriculum significantly improved students' critical thinking (Cohen's d=2.3) and ethical sensitivity.
- Utilizing the high-profile "Liping Chen–PD-1/PD-L1 Nobel priority controversy" as a structured academic controversy (SAC) case enabled students to master key immuno-oncology concepts without requiring additional curricular time.
- The scalable SAC model provides an efficient, evidence-based pathway to cultivate ethically-minded physician-scientists within the constraints of crowded medical curricula.

1 INTRODUCTION

Dual-process theory holds that within eight-year medical programs, students need to develop the ability to "evaluate the weight of the evidence". In practice, we have come to understand that Medical Immunology is inherently suited to developing this capacity [1-3]. Compared to other medical disciplines, immunology more frequently intersects and integrates with many fields [4, 5]. For example, the autoimmune destruction of pancreatic β-cells in type 1 diabetes demonstrates an immunology-endocrinology connection; the blockade of neural signals by anti-acetylcholine-receptor antibodies in myasthenia gravis illustrates nervous system-immunology interactions; and the engulfment of low-density lipoprotein by macrophages to form foam cells in atherosclerosis links the cardiovascular and immunological systems. Consequently, the cases encountered in a medical immunology curriculum require students to analyze and explain phenomena with multidisciplinary, multiangular, and multifaceted views. In such a complex mental structure, students must learn how to think.

Incorporating debate-based pedagogy into Medical Immunology seminars has been shown to be an effective medium for accomplishing educational goals [6, 7]. According to the revised Bloom's Taxonomy, an effective lesson starts with knowledge, through analysis, and culminates in evaluation [8, 9]. At the knowledge level, students must master the research areas and particular achievements of Lieping Chen's team—namely, the PD-1/PD-L1 pathway—which reinforces their recall of co-inhibitory molecules found in the T-lymphocyte unit. At the analysis level, students study on their own the Nobel Prize-winning work of James Allison and Tasuku Honjo, comparing the different paths taken regarding the CTLA-4 and PD-1 targets. Finally, at the evaluation level, an in-class debate allows students to construct and critique competing arguments, thereby developing their organizational skills and verbal agility. In summary, debate-centered instruction confers significant benefits by reinforcing theoretical knowledge and fostering the development of critical thinking skills in medical students.

The eight-year medical program is a combined "bachelor's-to-doctorate" track that confers a doctoral degree upon completion [10]. It aims to train physicians with profound medical knowledge and robust clinical and research capabilities. Consequently, students are required to develop a foundational understanding across various medical disciplines [11]. In comparison to shorter programs, this extended format strongly fosters independent learning. Students have to figure out when to study on their own, make a plan that suits them, and keep improving their own methods [12]. Over eight years of hard work and practice, they get used to working with classmates who have different kinds of expertise and personalities. This integration of everyone's best skills for collective improvement lays the foundation for effective teamwork in future hospital settings [13].

2 CURRICULUM DESIGN

2.1 Overall framework

2.1.1 Positioning and distinctive features

The study included 48 third-year students (26 females and 22 males) from the eight-year medical program. Among them, 11 (22.9%) had prior formal debate experience (e.g. university-level debate competitions), and 37 (77.1%) had no structured debate training. The debate-centered module, based on the "Lieping Chen Nobel Prize Controversy", serves as a signature case within the eight-year medical immunology curriculum. Focusing on the main arguments—concerning who made the first discovery of PD-1/PD-L1, the omission of a Nobel prize, and the unequal contributions of the involved teams—it brings together four different viewpoints: basic science, clinical application, ethical issues, and social influence.

2.1.2 Objective architecture

• Knowledge dimension: Mastery of the molecular mechanism of the PD-1/PD-L1 pathway and the principle of tumor immunotherapy.

• Competency dimension: Training in critical reading, evidence synthesis, academic debate, and ethical reasoning in research.

• Affective dimension: Fostering the development of a scientific ethos characterized by openness, truth-seeking, pluralism, and an appreciation for diverse contributions.

2.1.3 Schedule allocation

A two-hour seminar will be held one week after the lectures on B lymphocytes and antibody-mediated humoral immunity are completed.

Module	Time/Duration	Key Content	Learning Activities & Tasks	Deliverables	Evaluation Focus
Pre-class Micro-preparation	T-7 to T-1 day (extracurricular, fragmented)	Excerpts from Lieping Chen's PNAS paper & 2 pages of the Nobel White Paper	Students learn via mobile app; post a 50-character "biggest doubt" in the forum and upvote others	Top 3 most-liked doubts generated by the system	Focus of doubt; number of upvotes
In-class Rule Launch	0′–5′ (face-to-face)	SAC debate rules; timing; role descriptions	One flowchart & 40-second countdown GIF from instructor	Students understand steps and roles	Immediate classroom compliance
Position Building	5′–15′ (small-group, face-to-face)	Random draw to form Pro/Con sides	Form two opposing teams	Students confirm their stance	Immediate classroom compliance
Formal Debate	15′–65′ (whole class, face-to-face)	Opening statements 2′ × 2	Timer displayed & real-time bullet-screen voting	Live performance rubrics & word clouds	Radar chart from critical-thinking scale
Lightning Reflection	65′–80′ (face-to-face & online)	Six 5-level items (3 CT & 3 ES)	Scan QR code → instant radar chart	Class radar chart & individual scores	≥90 % completion; upward trend in radar
Post-class Micro-extension	T+1 to T+7 days (extracurricular)	60-second voice answer: "How to improve the Nobel selection?"	Upload voice on Rain Classroom & peer upvotes	"Golden-quote" posts with ≥5 upvotes	Originality of voice; number of upvotes

Figure 1. Overview of the debate-based teaching module design and assessment framework based on the "Lieping Chen Nobel Prize Controversy". PNAS, Proceedings of the National Academy of Sciences of the United States of America; SAC, Structured Academic Controversy; CT, Critical Thinking; ES, Ethical Sensitivity; GIF, Graphics Interchange Format; QR, Quick Response.

2.1.4 Pedagogical logic

The module follows an "online-offline-online" closed-loop structure, consisting of preparatory online learning, face-to-face debate, and online consolidation. The model's main framework is the SAC, which incorporates the strengths of Problem-Based Learning (PBL) and case-based instruction [14].

2.2 Breakdown of teaching modules

A comprehensive table details the structure and evaluation criteria of the debate-based teaching module centered on the Liping Chen Nobel Prize academic controversy (**Figure 1**). This figure shows the entire instructional process, from pre-class preparation to post-class consolidation. It specifies the timeline, core content, intended student learning activities, and expected learning outcomes. It acts as a practical guide for educators to carry out the module successfully and as a benchmark for evaluating its impact.

3 TEACHING IMPLEMENTATION

3.1 Pre-class phase: 7-day micro-preparation

The teacher releases a "Minimalist Learning Kit" via the Rain Classroom platform seven days before the discussion session. The kit contains four components: the full 2012 *Proceedings of the National Academy of Sciences of the United States of America* (PNAS) paper by Lieping Chen's team, a two-page excerpt from the Nobel Committee's 2023 white paper, a three-minute animation explaining the PD-1/PD-L1 mechanism, and a pre-filled "evidence quick-sheet" highlighting key chronological milestones. Students access the kit on their mobile devices, mark it as reviewed, and post their "biggest doubt" (under 50 characters) on the course forum. The system automatically tallies upvotes, and the top three most popular questions are identified for in-class focus.

3.2 In-class phase: 2-hour "debate sprint"

3.2.1 Time flow

The two-hour session (40 minutes × 2) is divided into four segments:

(1) 5-minute rule briefing: Using a single SAC flowchart, the instructor gives a quick briefing on the "position–rebuttal–switch" protocol.

(2) 10-minute position building: Students are randomly assigned to two groups—one supporting the Nobel Committee's decision and the other questioning Chen's omission. Each group drags key evidence from the shared "evidence quick-sheet" onto a Padlet to create a one-page A3 poster, projected on the classroom screen.

(3) 50-minute high-intensity debate: This phase includes opening statements (2 min) cross-examination (2 min), jury follow-up (5 min), position switch and rebuttal (2 min). Timing is displayed on screen, while the teacher simultaneously marks logical gaps on a digital whiteboard.

(4) 15-minute lightning reflection: Students scan a QR code to fill out a 5-level scale (assessing critical thinking and ethical sensitivity); the system instantly generates a class "cloud radar" chart. The instructor spends 2 minutes highlighting key takeaways, followed by a 13-minute open session where volunteers share their most significant insight of the day.

3.2.2 Space and staffing

A smart classroom with 48 fixed seats is used. Tables and chairs are arranged in a circular formation to create six round tables,

facilitating intra-group discussion while maintaining full visual contact among all participants. The teaching group is made up of three members: a primary instructor responsible for overall coordination and real-time feedback, one teaching assistant managing on-site timing, and another maintaining the Padlet board and rating-scale backend. Additionally, a technician from the Media Center supports audio-visual and networking functions.

3.2.3 Contingency plans

A three-tier contingency plan is in place: (1) Network failure: If the network goes down, the technician switches to a 5G hotspot within two minutes, while offline timers and paper rating scales are deployed. (2) Emotional escalation: If emotions are heightened during the debate, the ethics officer from the relevant small group raises a yellow card, triggering an immediate "30-second cool-down" pause. At this point, the lead instructor intervenes to facilitate. (3) Time overrun: If a speaker goes over their allotted time, an automatic buzzer sounds. The microphone is cut off after a 15-second grace period to keep the session on schedule.

3.3 Post-class phase: 24-hour "micro-reflection"

3.3.1 Micro-reflection and peer review

Each student has 24 hours to submit a 60-second voice message in response to the prompt: "if I were on the Nobel Committee, how would I improve the selection rules?" The system automatically transcribes the audio for peer review. Any statement receiving five or more "likes" is highlighted as a "golden sentence".

3.3.2 Data archiving

Class radar chart, Padlet posters, full debate video, and transcribed 60-second reflections are bundled and uploaded to the course cloud drive, which serves as the main data for subsequent teaching evaluation and educational research.

This process completes the integrated "pre-class, in-class, and post-class" instructional loop: students spend little time outside of class to prepare, engage in 80 minutes of high-density intellectual combat in the classroom, and complete a rapid reflection within 24 hours. This cycle not only generates reusable teaching and research data but also ensures a smooth fit with the preceding B-lymphocyte theory sessions and achieves the higher-order objectives of debate-based learning.

4 TEACHING EVALUATION

For scoring and evaluation, it is necessary to consider not only the performance of students in class but also their ability to transfer acquired knowledge to new immunological problems [15, 16]. Four aspects of evaluation, including student self-assessment, peer assessment, instructor assessment, and evaluation by two randomly selected student judges, should all be taken into account [17].

The rubric is divided into four domains:

- Knowledge mastery (30%): Gauges how thoroughly students grasp the PD-1/PD-L1 mechanism and the Nobel Prize controversy background.

- Critical thinking (30%): Assesses evidence screening, logical chains, and rebuttal strength.

- Ethical sensitivity (20%): Examines whether students recognize the tensions among scientific credit, public interest, and media narratives.

- Team collaboration & communication (20%): Measures division of labor, time control, and linguistic precision.

Temporally, evaluation spans the pre-class, in-class, and post-class phases. Ten percent of the grade comes from AI analysis of reading notes, assessing keyword density and depth of questioning. Sixty percent is derived from an eight-item, two-level rubric scored in real time by the instructor and student jury [18]. The remaining thirty percent is assigned to the 60-second post-debate voice reflection, first processed by AI for sentiment polarity and then weighted by peer upvotes.

To ensure consistent teaching methods across the cohort and uphold teaching quality, all eight-year program students participated in the debate-based instruction for this course. While this study lacked a control group using traditional seminar methods, the teaching reform yielded better results than the previous year's original teaching model. Mean critical-thinking score rose from 3.2 to 4.1 ($p<0.01$), and mean ethical-sensitivity score from 3.4 to 3.8 ($p<0.05$). Statistical analyses were performed using paired t-tests on the sample (N=48), after normality was confirmed via the Shapiro-Wilk test ($p>0.05$ for all scales). Effect sizes were Cohen's d=2.3 (critical thinking, large effect) and Cohen's d=1.1 (ethical sensitivity, medium effect), alongside Pearson's r=0.72 for the correlation between pre- and post-critical thinking scores. The 95% confidence intervals were 0.62-1.18 for critical thinking and 0.21-0.79 for ethical sensitivity.

When planning the next iteration, the instructor can refer directly to the class radar chart: if the mean ethical-sensitivity score falls below 3.5, the new case will foreground conflicts related to research integrity; if the mean logical-chain score exceeds 4.5, the complexity of the case will be increased. Thus, evaluation ceases to be merely an endpoint and becomes the starting point for the next cycle of instructional refinement.

5 TEACHING REFLECTION

After the session, the teaching team reflected on the lesson centered on the "Lieping Chen Nobel Prize Controversy", with a focus on the student debaters. Reflection was informed by multiple sources: classroom video analysis, student questionnaire results, peer observation notes, and feedback from an expert pane. The overall course reached its intended objectives, provided a positive student experience, and established a foundation for continuous improvement. However, some areas were identified as needing adjustment.

Most students correctly applied the core concepts of the PD-1/PD-L1 pathway and connected them to B-cell humoral immunity during the debate. But some relied excessively on the conclusions provided in the evidence quick-sheets without checking the experiments or statistics. To fix this, a 90-second "Methodological Zoom-In" micro-video will be added next time, focusing on the flow cytometry gating strategy from Lieping Chen's 2012 PNAS paper to help students practice their critical reading skills ahead of time [19].

Ethical sensitivity cultivation surpassed expectations, with a class mean score of 3.8 on the lightning reflection scale—an increase of 0.4 compared to a previous comparable course. Engagement with the scientific controversy was pivotal to their development of a professional identity, helping them connect ethical reasoning to their future roles as physician-scientists. They talked about how to balance acknowledging individual credit with appreciating collaborative scientific effort, an important skill for both research teamwork and hospital-based clinical decision-making. However, text analysis revealed that several students conflated "order of contribution" with "order of value", using phrases such as "just follow-up work". In order to fix this problem, the team plans to add a "contribution ladder" visualization to the next lessons. Students will categorize each scientist's work as discovery, validation, or translation, and consider how to apply principles of equitable authorship—translating abstract ethical concepts into actionable research governance practices.

One-month follow-up data revealed that 83% of students correctly used the PD-1/PD-L1 case to answer a question about antibody diversity, but only 47% could similarly apply their critical framework to novel targets such as CD47 or LAG-3. To enhance cross-case transfer, a five-minute online micro-debate will be added after the discussion of each subsequent chapter. This intervention aims to train students to expand their critical framework from a single case study toward a adaptable, lifelong thinking skill.

This study was conducted at one institution offering an elite eight-year program. Future adaptations may include implementation in 5-year medical tracks, where pre-class preparation might be reduced to three to four days and debate rules simplified to emphasize the synthesis of core evidence. For international settings, the model could incorporate region-specific scientific controversies and offer multilingual debate options, while retaining the "minimal preparation–high intensity interaction" structure to accommodate varying curriculum densities.

Observer bias is a limitation, as instructors also served as evaluators. Mitigation measures planned for future iterations include using external raters (e.g., faculty from other departments) and implementing blinded scoring (by anonymizing student identifiers in debate recordings and reflections prior to assessment).

6 CONCLUSION

This study shows that integrating a structured, debate-based module centered on the Lieping Chen Nobel Prize priority controversy into an existing eight-year Medical Immunology curriculum can enhance conceptual understanding, critical thinking, and ethical sensitivity, without requiring additional curricular time. Through a "minimal preparation–high-intensity debate–instant feedback" design, we found measurable improvements in knowledge transfer, quality of argumentation, and student engagement, while generating reusable teaching and research data. This model can be easily adapted to other controversial discoveries, providing a practical path for training evidence-based, ethically minded physician-scientists. Continuous refinement—such as adding micro-level methodological primers and cross-case transfer portfolios—are expected to further strengthen the long-term impact of this instructional approach.

DECLARATIONS

Author contributions

Liyuan Zhao and Yijie Tao drafted the main manuscript text. Sheng Xu and Min Zhang reviewed and revised the manuscript. All authors contributed to the article, reviewed the final version, and approved the submitted version.

Funding

This work was supported by the 2025 Teaching Achievement Cultivation and Development Project of the Naval Medical University (JPY2025B29).

Data availability

Not applicable.

Ethics approval and consent to participate

This study was approved by the Ethics Committee of the Naval Medical University (Approval Number: NMUMREC-2024-02002).

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

Not applicable.

REFERENCES

[1] Liu M, Zhang L. A two-pronged approach: reflections on promoting exercise behavior based on dual-process theory. J Phys Educ. 2024;31(06):64-70. https://doi.org/10.16237/j.cnki.cn44-1404/g8.20240927.005

[2] Cai J, Zhou J. Exploring a training model for high-level, innovative and versatile medical talents: A case study of the eight-year clinical medicine program. China Contin Med Educ. 2025;17(01):23-27.

[3] Lu Y, Pan Y. Reform and exploration of the "4+4" clinical medicine talent training model. Basic Clin Med. 2025;45(04):551-555. https://doi.org/10.16352/j.issn.1001-6325.2025.04.0551

[4] Wei L, Jiang Y, Tang H, Qin C, Haung G, Zeng Y. Innovation and practice of integrating "double innovation" into the "medical immunology" curriculum under the background of new medicine. Technol Wind. 2025;(09):81-83,150. https://doi.org/10.19392/j.cnki.1671-7341.202509027

[5] Mei J, Zhang D, Chen H. Practice and exploration of interdisciplinary integration of innovation and entrepreneurship education in medical immunology. Basic Med Educ. 2025;27(3):196-201. https://doi.org/10.13754/j.issn2095-1450.2025.03.02

[6] Cheng S, Meng Y, Zhao X. Preliminary exploration of debate-based teaching in eight-year TCM classrooms. Chin J Mod Distance Educ Tradit Chin Med. 2024;22(20):14-17.

[7] Feng J. Practice and rules of debate-based teaching in university classrooms. Chin Legal Educ Res. 2021;(04):113-129.

[8] Niu B, Ye P, Li Z, Luo S, Wei Y. Analysis of students' basic thinking patterns in PBL teaching based on Bloom-Turner questioning model. Technol Wind. 2025;(20):26-28. https://doi.org/10.19392/j.cnki.1671-7341.202520009

[9] Ren H, Liu X. A review of the application of Bloom's taxonomy to university curriculum design. Gov Dev. 2024;44-51.

[10] Xie Z, Wang W. Comparison and analysis of the eight-year clinical medicine curriculum system in eight domestic medical schools. China Higher Med Educ. 2025;(5):47-50. https://doi.org/10.3969/j.issn.1002-1701.2025.05.017

[11] Ma J, Shao L. Post-graduation development of eight-year clinical medicine students under the background of new medicine. J Sichuan Univ (Med Edit). 2025;56(02):596-602.

[12] Fang C, Li M, Liu L, Wang D, Xie A. Student evaluation and analysis of the quality of eight-year clinical medicine education. Med Philos. 2024;45(07):70-75.

[13] Chen X, Hu Y, Su J. Exploration and practice of scientific research training model for eight-year clinical medicine students. China Higher Med Educ. 2024;(9):24-26. https://doi.org/10.3969/j.issn.1002-1701.2024.09.009

[14] Rao Y. Discussion on the application of multiple teaching methods in medical courses. Course Educ Res. 2019;245.

[15] Jiang H. Analysis of the practical effect of college teachers' classroom teaching evaluation mechanism. J Jiamusi Vocat Coll. 2025;41(9):139-141.

[16] Zheng Q, Wsng S, Zhang J, Tai H, Luo X, Zhng J, et al. Differentiated status of medical teaching quality evaluation system and the construction path of core indicator sets. Chin J Med Guide. 2025;27(07):721-727.

[17] Suo C, Zhao Q, Wsng W, Chen X, Zhang T, Zhao G. Impact of teaching reform on the scale and evaluation of preventive medicine courses. China Contin Med Educ. 2020;12(14):9-12.

[18] Zhang J, Ren Y, Zhang J. Research and design of curriculum ideological and political evaluation based on Rubrics scale. Curric Ideological Political Teach Res. 2024;(1):80-97.

[19] Xie L, Yang M. Cultivation of critical reading awareness of medical students under the ESP framework. J Lang (Foreign Lang Educ). 2014;(02):97-100,142.