Received: 21 Jul. 2026 | Revised: 29 Jul. 2026 | Accepted: 12 Aug. 2026 | Published online: 25 Sep. 2026
RESEARCH ARTICLE

Contemporary Education and Teaching Research
2026, Vol. 7 (9)256-268
DOI: 10.61360/BoniCETR262020510901

A Questionnaire Research on Big Data Application and Compliance as General Courses in Medical Education

Jian Guan[1,*†], Tao Li[2†], Huanhuan Wang[3], Xiaolin Yang[4], Hua Peng[5], Yan Zhu[6], Cheng Feng[1], Tao Yang[1] & Lijuan Yu[1]

[1]National Population and Health Scientific Data Center (Clinical Medicine), Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College , Beijing, P.R. China
[2]School of Health Policy and Management, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, P.R. China
[3]School of Medical Information and Engineering, Xuzhou Medical University, Anhui, P.R. China
[4]Institute of Basic Medical Sciences & School of Basic Medicine, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, P.R. China
[5]Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, P.R. China
[6]Institute of Information on Traditional Chinese Medicine, China Academy of Chinese Medical Sciences, Beijing, P.R. China

Abstract: The status of data methodology and compliance courses in medical education in China was investigated to develop a curriculum design accommodating research paradigms and clinical practices transformed by big data and artificial intelligence. A total of 247, 241, and 233 valid questionnaires from medical professionals, undergraduates, and graduates from representative institutes were analyzed. Of the medical professionals, 25.51% (63/247) indicated that such courses were designed for undergraduates (73.02%, 46/247) and graduates (79.37%, 50/247). Systematic courses on medical data application and compliance were provided for 17.01% (41/241) of undergraduates and 56.65% (132/233)of graduates. More methodological courses were provided to undergraduates, accounting for 47.3% (114/241) of bioinformatics courses for basic medicine, 41.08% (99/241) of data mining courses for clinical medicine, and 86.27% (201/233) of graduates. However, responses confirmed that many respondents still lacked fundamental knowledge regarding data compliance. Totally, 97.57 (241/247) professionals and 97.47% (462/474) students believed that compliance courses to be more or equally important to methodology. The necessity of this course was discussed from the perspective of the purpose of general education and the acceptance of data-driven research by professionals and graduates. Finally, general courses for undergraduates were designed based on the preferences of the respondents and completed as preliminary practice.

Keywords: general medical education, graduate education, big data, artificial intelligence, methodology ,compliance, laws and ethics

1. Introduction

Big data and artificial intelligence (AI) are inseparable, and their applications are evolving medical development at an unprecedented pace, providing a technological foundation for the collection, sharing, and reuse of health data, with the potential to fundamentally transform medical research, clinical practice, and medical education (Guan J, 2025; McGrath SP et al, 2026). Big data and AI have transformed the paradigm of medical research, facilitating the shift from knowledge-based and experience-driven decision-making to data-driven clinical decision-making. However, with the rise of complex algorithms, new challenges have emerged, including algorithmic bias perpetuating health disparities, opacity undermining clinical trust, environmental sustainability concerns, and unresolved questions of privacy, security, data ownership, and interoperability (Raghupathi W et al., 2026). In addition, big data and AI have transformed teaching models and methodologies (Zhou et al., 2022; Zhang et al., 2022). Following the revolution triggered by breakthroughs in big data and AI technology, policy-makers have attempted to formulate strategies and policies regarding how to incorporate AI and emerging technologies into primary, secondary, and tertiary education (Pedró et al., 2019). China has enacted a series of laws and regulations to standardize data processing, security, and personal information protection. Furthermore, personal

Corresponding Author: Jian Guan
National Population and Health Scientific Data Center (Clinical Medicine), Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College , Beijing, P.R. China

information subjects (PISs) can be protected as research participants under the Helsinki Declaration (2024 version). On January 20, 2023, the Central Committee of the Communist Party of China and the State Council issued the landmark "Opinions on Establishing a Data-Based Institutional Framework to Better Leverage the Role of Data as a Key Factor." On April 10, 2026, the Ministry of Education convened a press conference to officially launch the "Artificial Intelligence + Education" Action Plan. The plan outlines a vision for the 15th Five-Year Plan period, detailing concerted efforts to advance and popularize AI education throughout society. From 2025, several medical universities established schools specializing in big data and AI, such as Peking University, Peking Union Medical College, and Sichuan University. The lawful and compliant application of medical data has become an imperative trend in medical education.

Traditional medical education content fails to meet contemporary educational demands. Big data methodology, governance, and management are particularly crucial. Governance and management involves addressing challenges and issues by regulating and standardizing the processes of data processing and utilization to ensure compliance with laws, regulations, and ethical principles.

To address this issue, we conducted surveys among students and professionals at medical schools and healthcare institutions to assess their understanding of big data applications in medicine, compliance considerations (legal and ethical aspects), and current curriculum frameworks. Building on these findings, we analyzed the challenges and demands of general medical education courses in the era of big data and AI. We believe that medical data application and governance should be incorporated into general medical education curricula based on respondents' preferences regarding course content and instructional approaches. We practiced undergraduate medical general education courses in 2026 and proved effective in complementing existing technical specialized courses through practical applications.

2.Basic Description

2.1 Questionnaires and statistics

An anonymous survey was developed in representative institutions from each administrative region of the Chinese mainland, with no fewer than two representative institutions per region, leveraging the convenience of the WeChat Mini-program for data collection.

We designed two questionnaires: one for medical professionals and the other for undergraduate and graduate students. Unless otherwise specified, the term "students" in this study refers to undergraduate students, graduate students, as defined by the survey questionnaire.Two questionnaires were used to collect data and perform statistical analysis, respectively. Both questionnaires addressed the key issues of the research purpose of this study. Considering the research requirements and time constraints, special questionnaires were designed for undergraduate and graduate students based on the same background questions.

To provide objective answers for the subsequent teaching system, when a participant did not promise a truthful answer, the process was automatically terminated (the participants did not know this design). In addition, each WeChat account could participate in the survey only once, including cases where the survey was terminated due to lack of commitment.

Leveraging our independently developed INSPIRE big data mining platform, we conducted comprehensive univariate statistical analysis.

2.2 Basic information of medical professionals

For medical professionals, out of the 251 invitations, we received 247 effective responses （98.4%）. Detail basic information are presented in table 1.

Most of the medical professionals were from healthcare institutions (65.99%, 163/247) and medical universities and colleges (25.91%, 64/247). Only 18 responses were from institutes (7.29%, 18/247). Professional title distribution: junior 15.1% (38/247), Middle-level title 33.9% (85/247), associate senior 27.5% (69/247), and senior 21.9% (55/247), others 1.6% (4/247). Of the respondents, 61.94% (153/247) were engaged in teaching work, but only 11.74% (29/247) were full-time teachers. Their teaching work included undergraduate courses (134, 54.25%) and graduate courses (68,27.53%). In addition, they participated in teaching the residency training program (67, 27.13%), guidance of basic course internships (31,12.55%), and supervision of graduate students (109, 44.13%).The detailed professional information of the medical staff is shown in figure 1A, 1C.

Table 1. Basic Information of Medical Professionals (Organization Type and Job and Teaching Work)

Organization type working	**Number**	**Percent**
From university and its colleges	64	25.91%
From healthcare institutions (hospitals)	163	65.99%
From research Institutions	18	7.29%%
From company or Enterprise	2	0.81%
Routine Job	**Number**	**Percent**
Clinical physician	100	40.49%
Full-time researcher	49	19.84%
Public health and management personnel	60	24.29%
Nursing	7	2.83%
Other job	31	12.55%
Professional title	**Number**	**Percent**
Primary title	38	15.38%
Middle-level title	85	34.41%
Associate Senior	69	27.94%
Senior title	55	22.27%
Teaching work		
Part-time teaching	124	50.2%
Full-time teaching	29	11.74%
No teaching	94	38.06%%
Teaching undergraduate courses	134	54.25%
Teaching graduate-level courses	68	27.53%
Participate in teaching activities for the residency training program	67	27.13%
Participate in the teaching and guidance of basic course internships	31	12.55%
Participate in clinical internship supervision	99	40.08%
Participate in residency training and clinical supervision	57	23.08%
Supervise graduate students	109	44.13%

**Total respondents: 247*

2.3 Students basic information

We received 474 valid responses (98.4%) out of the 482 invited students. Among them, undergraduate students accounted for 50.84% (241/474), and graduate students accounted for 49.16% (233/474). Among them, 63 (26.14%) came from medical universities and 178 (73.68%) from comprehensive universities.Detail basic information and professional distribution are presented in table 2, and figure 1B, 1D.

Table2. Overview of Basic Information of Students

Organization type working	**Number**	**Percent**
Under graduate	241	50.84%
graduate	233	49.16%
Under graduates	**Number**	**Percent**
For 5-year undergraduate programs,duration of basic course studies ranges from grade 1-3	101	41.91%
For 5-year undergraduate program, duration of practical internship semester, grade4-5(clinical medicine)	33	13.69%
For 5-year undergraduate program, duration of practical internship semester, grade4-5(Other specialties)	7	2.9%
For 8-year undergraduate, duration of basic courses phase	25	10.37%
For 8-year undergraduate, duration of clinical courses phase	9	3.73%
For 4-year undergraduate, Nursing	66	27.39%
Graduates	**Number**	**Percent**
Academic & research-oriented	86	36.91%
Professional & clinical-oriented	144	61.8%
Others	3	1.29%

**Total respondents :474*

3. Current Status of Medical Data-related Curriculum Design

3.1 Answer of the medical professionals

Regarding medical data application compliance courses, which refer to specialized courses on data law, data security, or data governance management, only 25.51% (63/247) of the medical professionals reported that they had initiated such courses. Of the respondents, 18.62% (46/247) and 20.24% (50/247) reported that their institutions have designed

courses for undergraduate and graduate students, respectively.

Regarding the course content of courses already offered, 39.68% (25/63) included systematic courses on medical data application and related legal regulation management, while 34.92% (22/63) consisted of basic medical data knowledge courses, and 7.94% (5/63) consisted solely of big data and applied methodology courses. Additionally, 17.46% (11/63) of the surveyed professionals reported that they did not know about the specific course content. Only 10.12% (25/247) of all surveyed medical professionals showed their institution had established systematic courses on medical data application and compliance.

3.2 Answer of the undergraduates

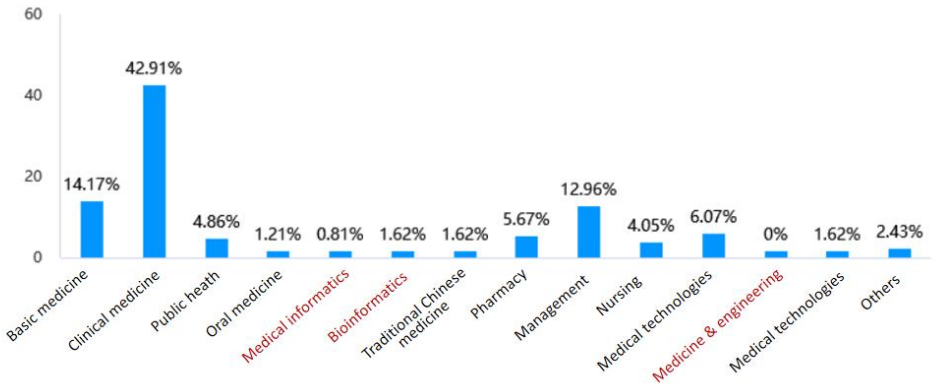
Among undergraduates, only 17.01% (41/241) believed that their universities provided systematic, specialized courses on medical data application and compliance with relevant laws and regulations; the vast majority remained unaware of such offerings. A notable 31.54% (76/ 241) had taken undergraduate courses covering fundamental knowledge of medical data, while a smaller segment,4.98% (12/241), had engaged with only courses on big data and methodology. Additionally, 14.52% (35/241) and 31.95% (77/241) answered that there were no related courses, and others were unaware of whether such courses were available.

Regarding courses involving data methodology, 47.3% (114/241) of respondents indicated that bioinformatics courses had been offered in basic medicine, while 41.08% (99/241) revealed that clinical medicine education had introduced data mining and analysis methodology courses.

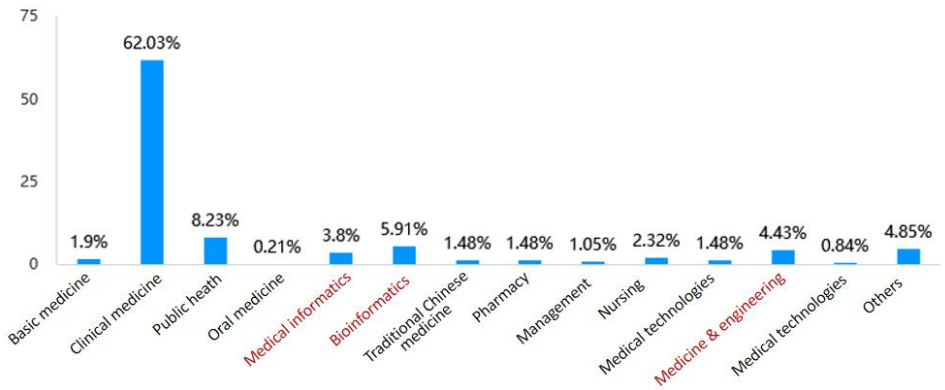
3.3 Answer of graduates

Among the graduate responses, 56.65% (132/233) of graduate students indicated that their universities provided systematic, specialized courses on medical data application and compliance with relevant laws and regulations. Regarding courses on medical data analysis methods, the course offers a diverse range of content, not limited to traditional statistics courses (86.27%,201/233) . Additionally, 37.77% (88/233) offered elective courses in bioinformatics and medical informatics. Only 4.29% (10/233) of the respondents indicated that no relevant courses were available.

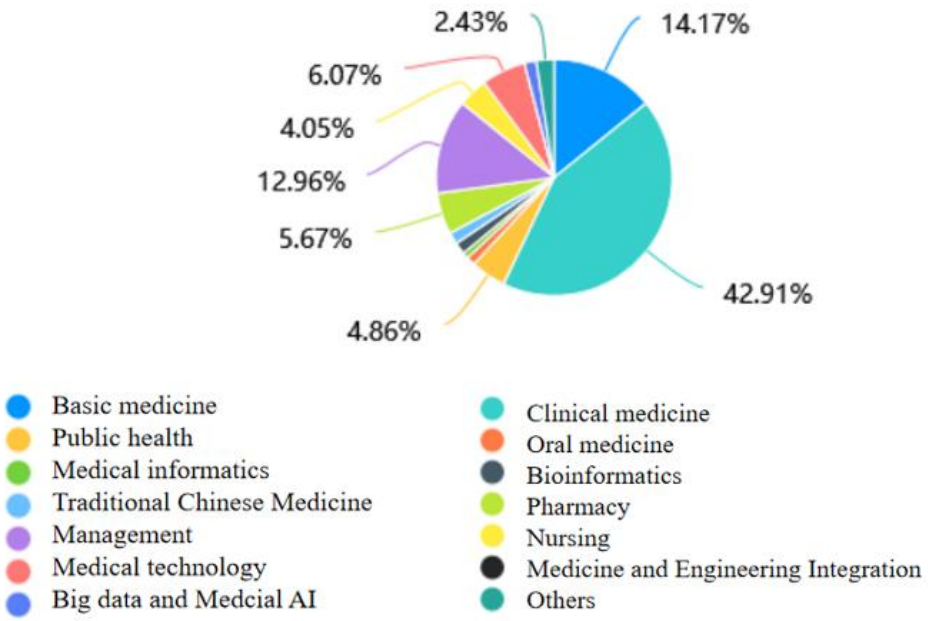
A



B



C



D

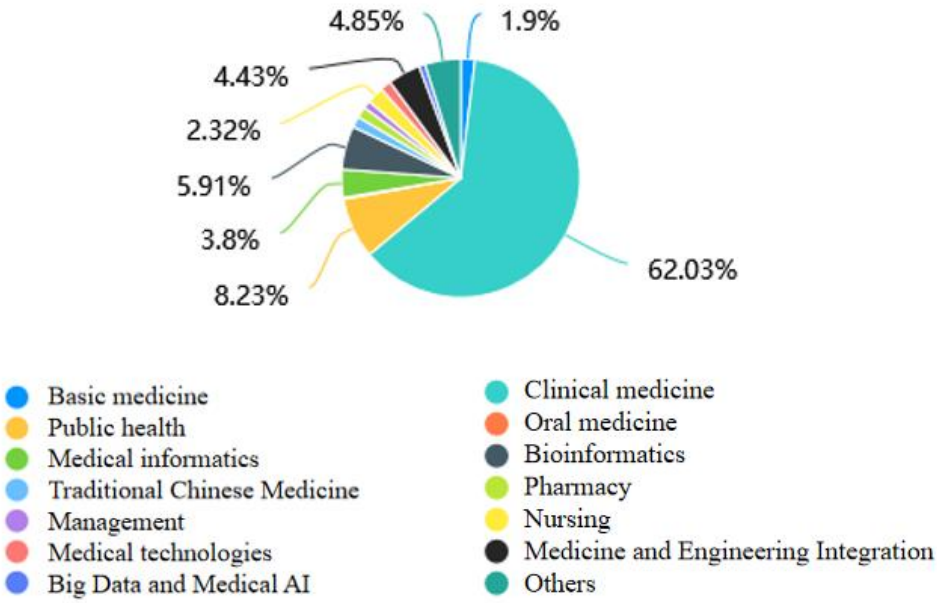


Figure 1. Professional Distribution of the Surveyed Medical Professionals and Students

**A: Bar chart showing the professional distribution of the medical staff*
**B: Bar chart showing the professional distribution of students*
**C: Pie chart showing the professional distribution of the medical staff*
**D: Pie chart showing the professional distribution of students*

The results indicate that regarding the establishment of systematic courses, the proportions reported by healthcare professionals (10.12%,25/247) and undergraduate students (17.01%,41/241) were relatively similar. Among graduate students, over 50% reported that systematic courses on data usage and compliance had already been established.

4. General Medical Education Has Teaching Requirements for Knowledge of Medical Data Application and Compliance

4.1 Data-driven research participated and considered by medical personnel and graduate students

With the rapid evolution of big data and AI, data-driven research has emerged as the dominant paradigm and primary methodology for researchers and graduate students.

4.1.1 Most medical professionals using data for research

The majority of participating medical professionals were clinical physicians, accounting for 40.49% (100/247), public health and management personnel accounting for 24.92% (60/247), full-time researchers made up 19.84% (49/247), and nursing action and "others" representing 2.83% (7/247) and 12.55% (31/247) of the total. Our survey revealed that 72.47% (179/247) of medical professionals collected and utilized data for their research or academic writing. Furthermore, 47.77% (118/247) accumulated, screened, and organized the data, while 44.94% (111/247) considered using the data for further research and writing endeavors.

It was clearly stated that 30.77% (76/247) of researchers did not consider solely applying data to conduct their studies, primarily due to objective reasons, including incompatibility with the research direction (23.68%,18/76), inability to obtain high-quality data (35.53%,27/76), and other constraints (40.79%, 31/76). Therefore, both the methodologies and compliance aspects of medical data applications are crucial in medical education.

4.1.2 More than half of graduate students participate in or utilize data for their research

The majority of participating graduate students were professional and clinical-focused, accounting for 61.8% (144/233); academic and research-oriented students made up 36.91% (86/233); and“others” represented 1.29% (3/233) of the total. Among these students, 82.83% (19/233) had participated in data collection tasks or utilized data to complete surveys or research papers, a proportion higher than that among professionals. Only 17.17% did not engage in such activities, which aligns with the fact that graduate programs include numerous courses covering data methodology.

4.1.3 Data methodologies are needed as general courses for medical education

16.19% (40/247) of the medical professionals and 24.89% (58/233) of graduate students chose not to consider utilizing existing data for their research or papers. The reasons primarily stemmed from three factors: among medical professionals, the most common reason was "uncertainty about specific research methods" (42.5%,17/40), followed by "inability to obtain high-quality data" and "incompatibility with research objectives" (37.5%,15/40; 20%,8/40); among graduate students, the least cited reason was "uncertainty about specific research methods" (22.41%,13/58), with other reasons being "inability to obtain high-quality data" and "incompatibility with research objectives" (43.10%,25/58; 34.48%, 20/58). This trend is related to students 'current academic disciplines and the mandatory methodology courses offered at the graduate level. These findings suggest that medical data methodologies should be included as fundamental components of general education. Additionally, they highlight how research methods, data quality, and research objectives impose constraints on faculty's use of existing data, warranting consideration in continuing education programs.

We believe that the core value of general education lies in laying a solid foundation for academic and professional development. These courses aim to cultivate well-rounded competencies and impart interdisciplinary skills to students. The results above indicate that data-driven research has been widely embraced and adopted by healthcare professionals and graduate students alike. Medical data application and compliance have become fundamental professional competencies, and medical education, research, and clinical practice rely heavily on big data and AI. Research capabilities and clinical practice are vital components of medical education. The ability to apply medical data in compliance with legal requirements is directly related to medical students' academic performance and future career development, whether in clinical practice or medical research, including the burgeoning field of AI in medicine. Consequently, the application of such data, including AI, alongside related compliance considerations, forms an indispensable cornerstone of modern medical education. In the era of big data and AI, the general education curriculum must evolve accordingly. Thus, general medical education has teaching requirements for knowledge of medical data application and compliance from the perspective of general education.

4.2 The respondents' comprehension status regarding the basic knowledge of medical data governance and management

Regarding evaluations of the concepts and understanding of data and its applications of the respondents, among healthcare professionals and students, we designed subjective self-assessment questions and objective multiple-choice survey items (covering fundamental concepts and comprehension of data and its applications).

According to the subjective evaluation results, more than half (52.63%,130/247) of the professionals do not understand the specific meaning of leveraging data as a key element; among them, 48.99% (121/247) have only a limited understanding, while 3.64% (9/247) have no understanding at all.

Similarly, 35.86% (170/474) of the total students directly selected "understood"; the proportions of those who were not quite understanding and those who completely did not understand were 57.17% (271/474) and 6.96% (33/474), respectively.

Responses to objective multiple-choice items related to data governance and management confirmed that many respondents lacked fundamental knowledge of data compliance. Regarding the sources of medical scientific data, the proportions of incorrect understanding among healthcare professionals and students were 32.79% (81/247) and 40.3% (191/474), respectively. For the statement that "medical data constitutes records of health-related information," the incorrect response rates among healthcare professionals and students were 22.27% (55/247) and 25.53% (121/474), respectively. Regarding the use of clinical medical data or other operational data for research or other purposes, the incorrect response rates were 17.81% (44/247) and 23.84% (113/474), respectively.

Furthermore, regarding multiple-choice questions concerning the laws, regulations, and provisions that must be complied with when applying medical data, 37.25% (92/247) of healthcare professionals and 27.85% (132/474) of students selected incorrect responses.

The results for this question reveal a significant gap between undergraduate and graduate students of basic knowledge of data application and governance; graduate students demonstrate a higher proportion of "understanding." For example, regarding the understanding of the role of data as production elements, 54.51% (127/233) of graduate students reported "understanding," compared with only 27.80% (67/241) of undergraduate students. This finding aligns with the observation that graduate students have a higher proportion of systematic data application skills and follow established curriculum requirements.

Simultaneously, this highlights the need to strengthen general medical education for undergraduate students in this field. Changes and challenges, such as research, industry, and policy-making of big data and AI in education, must be overcome to integrate big data and AI into educational practice (Luan et al., 2020). It is known that the application of big medical data and AI faces significant social and ethical challenges. One of these is personal rights, such as privacy protection and informed consent, for medical data, which are sensitive personal health information. Governance and management involve addressing challenges and issues by regulating and standardizing data processing and utilization to ensure compliance with laws, regulations, and ethical principles.

4.3 Medical professionals and students recognized the importance of both courses

The questionnaire showed that 82.19% (203/247) professionals considered both aspects equally important, 15.38% (38/247) deemed data compliance more critical, while only 2.43% (6/247) regarded data mining methods as more significant.

Notably, the proportion valuing both equally was highest among mid-level professionals at 83.5% (71/85). This group also constituted 34.98% (71/203) of all respondents who selected a balanced view. When prioritizing data compliance, associate senior professionals showed the highest rate at 18.84% (13/69), closely followed by full senior professionals at 18.19% (10/55). Interestingly, no full senior professionals indicated a greater focus on data mining methods.

Regarding all students (undergraduates and graduates), 83.54% (396/474) considered both courses equally important, 13.92% (66/474) deemed "data compliance more critical", and only 2.53% (12/474) regarded “data mining methods as more significant". Of the undergraduates, 85.06% (205/241), 12.03% (29/241), and 2.9% (7/241) chose "both methodology and complicance equally important", "compliance is more critical" and "methodology is more critica", respectively. Of the graduates, 81.97% (191/233), 15.88% (37/233), and 2.15% (5/233) chose "both methodology and compliance equally important", "compliance is more critical" and "methodology is more critical" respectively.

In total, 97.57 (241/247) of professionals and 97.47% (462/474) believed that compliance courses were more or equally important to methodology. Table 3 presents a detailed breakdown of the choice of courses by professional title, undergraduates, and graduates.

These results demonstrate that both professionals and students recognize the critical importance of data compliance and data mining

methodologies for medical education and future work, underscoring the necessity and feasibility of this course design.

Table 3. Perspectives on the Importance of Data Processing Methods and Compliance for Developing Curriculum as General Courses

Academictitle	Big data-related methodology course & Big-datarelated compliance courses			
	Both are equally important (N)	Data compliance is more important (N)	Data mining methods are more important (N)	Total(N)
Primary title	32	5	1	38
mid-level Professional title	71	10	4	85
Associate professional title	55	13	1	69
Senior professional title	45	10	0	55
Total	**203**	**38**	**6**	**247**
Graduate student	205	29	7	241
graduate student	191	37	5	233
Total	**396**	**66**	**12**	**474**

**N: number*

5. General Medical Courses on Medical Data Applications and Compliance Built Based on Preferences of Medical Professionals and Students

5.1 Course and assessment method preferences

A course preference in the questionnaire was designed tailored to the distinct educational objectives of undergraduate, graduate.

5.1.1 Design the curriculum from the perspective of professionals

The combination of course requirements and assessment methods applies to the existing courses: percentage of required courses was 61.90% (39/63), combination of required course, exam-based course, closed-book test account for 33.33% (21/63), followed by elective courses; practical case analysis open-book test 19.05% (12/63), elective courses; assessment courses, closed-book test 12.7% (8/63). This may differ from the arrangements for undergraduate and graduate students or be specifically tailored to programs in big data and AI.

Regarding medical personnel who had not yet enrolled in the course, their preferences for future course offerings were as follows: 115 professionals selected compulsory courses, accounting for 62.5% of valid responses (115/184). Unlike existing courses, 66.96% (77/115) preferred open-book examinations. Among those who chose elective courses, 37.5% (69/184) supported this approach, with 63.77% (44/69) specifically advocating for open-book exams. Overall, regardless of enrollment requirements, 65.76% (121/184) of students preferred open-book examinations. These results clearly demonstrate strong recognition among professionals of the course's necessity, while also highlighting that open-book exams serve purposes beyond mere assessment.

We preliminarily investigated the impact of professional titles on course perceptions (Figure 2). The study revealed that individuals with different professional titles may have varying perceptions of medical data application and compliance courses: those with senior titles (full professor or associate professor) place greater emphasis on the practicality and cutting-edge nature of the courses, favoring formats that include practical case analyses; those with intermediate titles prioritize systematicness and standardization, showing higher acceptance of assessment methods; whereas those with junior titles prefer more foundational courses focused on building a comprehensive knowledge framework.

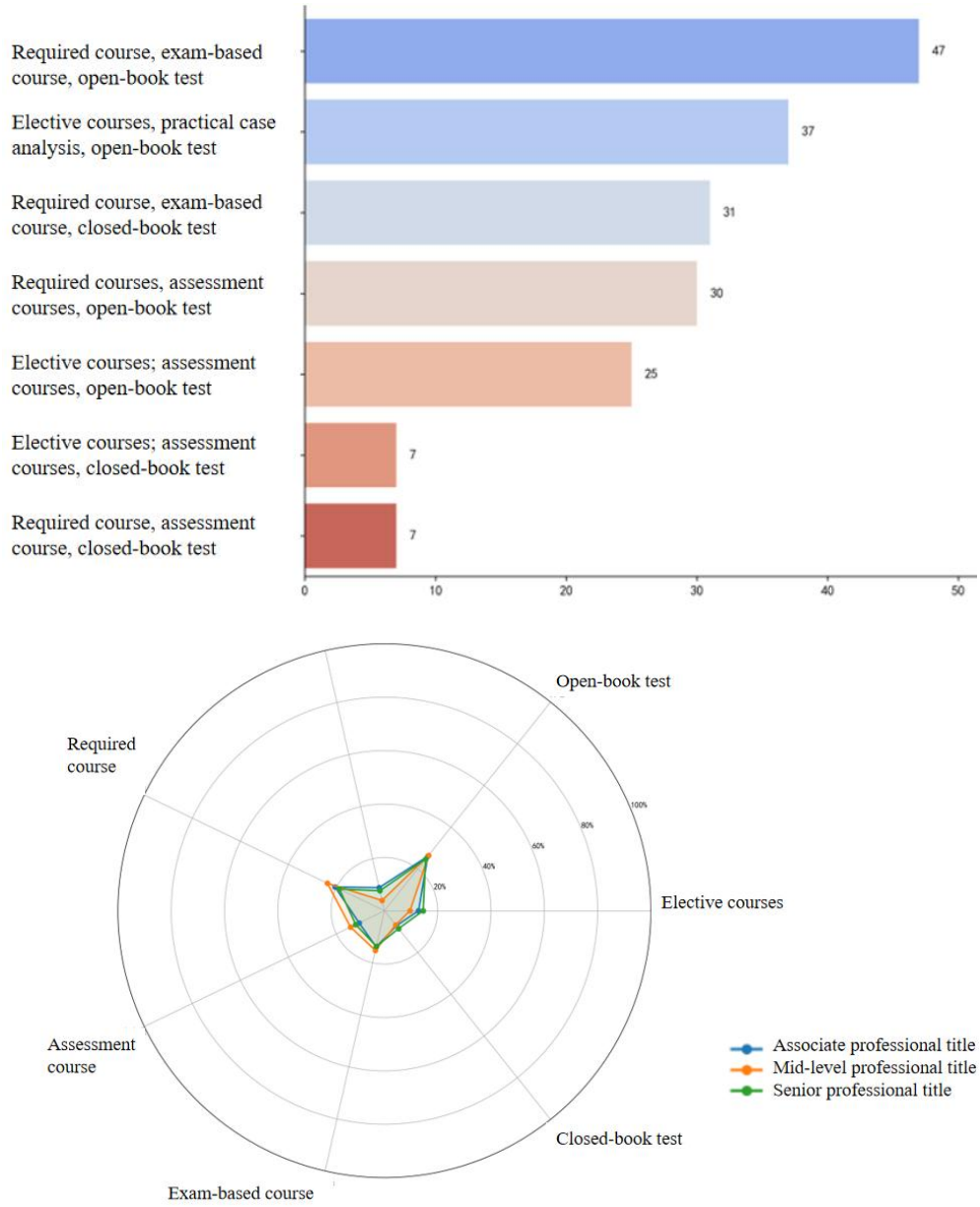


Figure 2. Preferences for course requirements and assessment methods across various professional titles. Up: bar chart; Down: radar chart

5.1.2 Student preferences

Students' preferences for courses and assessment methods reflect their learning needs and desire to reduce exam-related pressure. Among undergraduates, 80.91% (195/241) preferred open-book exams, while a higher proportion 62.66% (151/241) opted for elective. Graduate students showed an even stronger preference for electives, with 83.69% (195/233) choosing them; additionally, 60.94% (142/233) selected assessment courses, and39.06% (91/233) chose open-book tests for practical case analyses. Overall, a high proportion of medical professionals, undergraduate and graduate students, and those enrolled in both compulsory and elective courses opted for open-book examinations. The summarized preferences regarding course requirements and assessment methods are shown in Table 3.

Table 4. Preferences of Course Requirements and Assessment Methods for General Courses on Medical Data Application and Compliance

Respondents	Combination of course requirements and assessment methods	Existing courses		Prefer choice for expected courses	
		N	Percent	N	Percent
Medical professionals	Required course, exam-based course, closed-book test	21	33.33%%	31	16.85%
	Required course, exam-based course, open-book test	8	12.7%	47	25.54%
	Required course, assessment course, closed-book test	5	7.94%	7	3.8%
	Required courses, assessment courses, open-book test	5	7.94%	30	16.3%
	Elective courses, assessment courses, open-book test	4	6.35%	7	3.8%
	Elective courses; assessment courses, closed-book test	8	12.7%	25	13.59%
	Elective courses; practical case analysis open-book test	12	19.05%	37	20.11%
	Total	**63**	**100%**	**184**	**100%**
Under graduates (Total 241)	Required course, assessment course, closed-book test		/	25	10.37%
	Elective courses, assessment courses, closed-book test		/	21	8.71%
	Required course, assessment course, open-book test		/	65	26.97
	Elective courses, assessment courses, open-book test		/	130	53.94%
	Total			**241**	**100%**
graduates	Required course, assessment course, closed-book test		/	38	16.31%
	Elective courses, assessment courses, closed-book test		/	17	7.3%
	Elective courses, assessment courses, open-book test		/	87	37.34%
	Elective courses; practical case analysis open-book test.		/	91	39.06%
	Total			**233**	**100%**

**N: number*

5.2 General course we designed for medical undergraduates based on preferences and AI development, and completed as preliminary practice

To improve teaching practices, we explored the design of courses for general medical education. The courses were based on the preferences of the respondents; however, AI development in medical fields along with big data applications were also considered.

Systemic courses are urgently needed in undergraduate teaching; thus, we designed general education courses for undergraduate students in medical schools who do not major in big data or AI as follows.

This course serves as a foundational professional course for undergraduates majoring in clinical medicine programs (including Traditional Chinese Medicine and Nursing), is designated as an elective (extension course), and falls under the general education component of the undergraduate medical education reform. The course comprises 30 credit hours and adopts a formative assessment model with an open-book examination.

Course objectives: This course aims to cultivate clinical medical professionals and researchers who can keep pace with the era of big data and AI; to enable medical students to gain a comprehensive understanding of fundamental knowledge in traditional medicine, as well as foundational legal and ethical knowledge of big data, AI, and their applications in the medical field, along with the corresponding legal and ethical requirements applicable to various practical scenarios; to equip medical students with an understanding of clinical medical responsibilities and medical research duties in modern medical contexts; to cultivate advanced undergraduate-level medical professionals aligned with the evolving landscape of modern medicine; and to prepare high-level medical professionals with the

knowledge base required for postgraduate education and interdisciplinary talent development.

Course consents: The course consists of three units: Fundamentals and Overview of Medical Data and Artificial Intelligence; Applications of Medical Data – Basic Requirements for Laws, Regulations, and Ethical Standards; and Medical Artificial Intelligence – Laws, Regulations, and Ethical Standards.

Course tasks: Upon completion of this course, students will be able to:

(1) Understanding the fundamental concepts of medical data and AI, how medical data serve as a key data element in clinical practice and medical research, and the basic application methods through which AI empowers medicine and research;

(2) Understanding the essence and unique characteristics of medical data, along with foundational knowledge of the sources, types, and features of medical big data;

(3) Familiar with the legal and regulatory framework governing the application of medical data, as well as the basic principles of data ethics;

(4) Understanding fundamental concepts, such as the definition and classification of AI, its advantages and challenges in the medical field, and the strengths and challenges associated with generative AI;

(5) Mastering the key legal provisions and ethical guidelines governing the application of medical data and AI; this course aims to encourage students to consciously enhance the effective utilization and reutilization of medical data in their clinical practice;

(6) Enhancing students 'data and AI literacy, as well as their legal and ethical awareness regarding medical data and AI;

(7) Increasing medical students' understanding of the functionalities of generative AI and fostering the development of their scientific integrity competencies in its educational and research applications.

In December 2025, we confirmed the teaching syllabus and course assessment methods with our partner institution, Shenyang Medical College. Between April and June 2026, we completed the course teaching plan and assessment procedures, thereby achieving the general education objectives for medical students. The core content of this course (Law and Ethics) comprises a total of 8 credit hours, which has also been incorporated into the curriculum plan of the partner institution, Hangzhou Medical College, and will continue to be refined and enhanced.

Big data applications are inextricably linked to AI; moreover, as complex algorithms, AI involves ethical and legal dimensions that intersect with data ethics. In recent years, big data in healthcare and medicine, evaluated based on shift from Hadoop-centric architectures to cloud computing platforms and GPU-accelerated AI, including large language models and the emerging paradigm of agentic AI (Raghupathi W et al., 2026). Future medical research and clinical practice (smart healthcare) will rely heavily on big data and AI. AI is rapidly reshaping medical research and healthcare, which takes the requirements of the competencies expected of graduating medical students changed; however, AI curricula and competency recommendations for general medical education remain fragmented (Hunt VM et al., 2026). In addition, AI competencies were proposed for undergraduate medical education from 564 eligible statements, resulting in a taxonomy comprising seven domains, including clinical application, AI Ethics, AI Law and Regulation (Hunt et al., 2026).

China's education authorities have recognized the educational trends related to data and AI. The core objective of "Artificial Intelligence + Education" Action Plan is to establish a comprehensive AI education system by 2030 that spans all educational stages and connects across disciplines, alongside a general education framework for society. We hope that this research can provide ideas for medical colleges and universities.

6. Conclusion

This study examined the current status of data methodology and compliance as general courses in medical education at universities and colleges in China, focusing on the needs and preferences of professionals and students (undergraduates and graduates) regarding the effects of big data and AI on medical research and services. The findings indicate that systematic courses for medical curricula were provided only in a few medical universities and colleges. Most were focused on data methodology, thus particularly lack of courses on compliance for data processing and application. However, medical professionals and graduate students have already shown interest in and have planned to utilize the data for conducting research and writing academic papers. The vast majority of medical professionals and students recognized the importance of this course.

In light of the general education courses required in the era of big data and AI, as well as the educational requirements under the "Artificial Intelligence+" initiative for relevant departments, data methodology and compliance knowledge are essential components of general medical education and should be incorporated into the general curriculum of medical education. Specific course arrangements can be determined based on

teaching objectives, academic discipline, and the preferences of medical professionals and students. This study also investigated and summarized the preferences of the respondents, which will effectively provide references to universities and colleges. As general education courses, they can be designated as either compulsory or elective based on student preferences and requirements, and are assessed through an open-book examination. Preliminary practical implementation has demonstrated that general education courses and examination formats tailored to students' preferences can effectively achieve general education objectives for medical students.

Conflict of Interest

The authors declare that they have no conflicts of interest to this work.

Acknowledgement

This research was supported by the Undergraduate Education and Teaching Reform Project of PUMC (2023zlgl059); National Science and Technology Infrastructure Center Operation Platform: National Population Health Science Data Sharing Service Platform Operation Funds.

Limitations of the Study

The convenience of selecting institutions and conducting surveys, coupled with sample size constraints, resulted in relatively low representation in certain regions. Regarding the existing curriculum design, courses for graduate students have been further refined; however, they do not include the same questions as those designed for medical professionals and undergraduate students. While general education courses for undergraduates have undergone preliminary validation through practical applications, graduate courses lack specific practical components.

References

Guan, J. (2026). Fundamental Overview of Medical Scientific Data Sharing. In: *Governance and Management of Medical Scientific Data Sharing and Application.* Springer, Singapore. https://doi.org/10.1007/978-981-95-2806-6_1

Guan, J. (2026). Governance and Management of Medical Artificial Intelligence. In: *Governance and Management of Medical Scientific Data Sharing and Application*. Springer, Singapore. https://doi.org/10.1007/978-981-95-2806-6_14

Hunt VM, Sprehe LK, de Lomba WC, Gameiro RR, Woite NL,Wade C G, et al. (2026). What the AI era doctor should know: a scoping review of proposed artificial intelligence competencies for medical education. *npj Digit. Med.* https://doi.org/10.1038/s41746-026-02761-9

Luan H, Geczy P, Lai H, Gobert J, Yang SJH, Ogata H, Baltes J, Guerra R, Li P, Tsai CC. (2020). Challenges and Future Directions of Big Data and Artificial Intelligence in Education. *Front Psychol.* 11:580820. doi: 10.3389/fpsyg.2020.580820. eCollection.

Lüdorf V, Meister S, Mainz A, Ehlers JP, Nitsche J, Busse TS. (2025). Developing a Concept on Ethical, Legal and Social Implications (ELSI) for Data Literacy in Health Professions: A Learning Objective-Based Approach. *Healthcare, 13,* 2108. https://doi.org/10.3390/healthcare13172108

McGrath SP, Kim KK, Johl K, Wang HB, Anderson N. (2026). AI-PACE: a framework for integrating AI into medical education. *npj Digit. Med.* https://doi.org/10.1038/s41746-026-02768-2.

Pedró F., Subosa M., Rivas A., Valverde P. (2019). *Artificial Intelligence in Education: Challenges and Opportunities for Sustainable Development.* Paris: UNESCO.

Raghupathi W, Raghupathi V. (2026). Big data in healthcare and medicine revisited: design and managerial challenges in the age of artificial intelligence. *Health Information Science and Systems, 14,* 38 . https://doi.org/10.1007/s13755-026-00433-2

Seth P, Hueppchen N, Miller SD, Rudzicz F, Ding J, Parakh K, Record JD. (2023). Data Science as a Core Competency in Undergraduate Medical Education in the Age of Artificial Intelligence in Health Care. *JMIR Medical Education, 9.* https://doi.org/10.2196/46344.

Zhang R. (2022). Digital Media Teaching and Effectiveness Evaluation Integrating Big Data and Artificial Intelligence. *Computational Intelligence and Neuroscience.* 2022:1217846. doi: 10.1155/2022/1217846. eCollection 2022.PMID: 36188689

Zhou X, Wu X. (2022). Teaching Mode Based on Educational Big Data Mining and Digital Twins. *Computational Intelligence and Neuroscience,* 2022:9071944. doi:

10.1155/2022/9071944. eCollection 2022.Retraction in: Comput Intell Neurosci. 2023 Aug 16;2023:9838956. doi: 10.1155/2023/9838956.

How to Cite: Guan, J., Li, T., Wang, H.H., Yang, X.L., Peng, H., Zhu, Y., Feng, C., Yang, T., & Yu, L.J.. (2026). A Questionnaire Research on Big Data Application and Compliance as General Courses in Medical Education. *Contemporary Education and Teaching Research, 07*(09), 256 − 268. https://doi.org/10.61360/BoniCETR262020510901