BMJ Open

Oncologic big-data platform in promoting digital competencies and professionalism in Chinese medical students: a cross-section study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2022-061015
Article Type:	Original research
Date Submitted by the Author:	12-Jan-2022
Complete List of Authors:	Liu, Jiahao; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Jiao, Xiaofei; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Zeng, Shaoqing; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Li, Huayi; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Jin, Ping; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Chi, Jianhua; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Liu, Xingyu; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Yu, Yang; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Ma, Guanchen; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Zhao, Yingjun; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Li, Ming; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Peng, Zikun; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Huo, Yabing; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Gao, Qing-Lei ; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology
Keywords:	MEDICAL EDUCATION & TRAINING, Gynaecological oncology < GYNAECOLOGY, Gynaecological oncology < ONCOLOGY, Health informatics < BIOTECHNOLOGY & BIOINFORMATICS

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Oncologic big-data platform in promoting digital competencies and professionalism in Chinese medical students: a cross-section study

Jiahao Liu[1,2#], Xiaofei Jiao[1,2#], Shaoqing Zeng[1,2], Huayi Li[1,2], Ping Jin[1,2], Jianhua Chi[1,2], Xingyu Liu[1,2], Yang Yu[1,2], Guanchen Ma[1,2], Yingjun Zhao[1,2], Ming Li[1,2], Zikun Peng[1,2], Yabing Huo[1,2], Qinglei Gao[1,2*]

[1] Department of Obstetrics and Gynecology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, Hubei, P. R. China.
[2] Cancer Biology Research Center (Key Laboratory of the Ministry of Education), Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, Hubei, P. R. China.

[#] These authors contribute equally.

* Correspondence to:
Prof. Qinglei Gao, Cancer Biology Research Center (Key Laboratory of the Ministry of Education), Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, 1095 Jiefang Ave, Wuhan 430030, Hubei, P. R. China. Tel: +86-27-83662681; Fax: +86-27-83662681; E-mail: qingleigao@hotmail.com.

Word count (excluding title page, abstract, references, figures and tables): 2,774

Abstract

Objectives: Advancements in big data technology are reshaping the health care system in China. However, medical students' attitude for big data technology, as well as their professionalism in the big data age were poorly investigated. This study aims to explore the role of medical big data in promoting the digital competencies and professionalism in Chinese medical students.

Design, setting and participants: This cross-sectional study was performed among 274 medical students who attended a workshop about medical big data conducted on July 8, 2021 in Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China. The workshop was based on the first national-wide multi-function gynecologic oncology medical big data platform in China, the National Union of Real-World Gynecologic Oncology Research & Patient Management Platform (NUWA platform).

Outcome measures: The knowledge and attitude towards big data technology and a Professionalism Mini-Evaluation Exercise (P-MEX) instrument were collected before and after the workshop.

Results: 274 students participated in this workshop and completed all surveys. Before the workshop, only 27% of them knew the detailed content of medical big data platforms and 64% knew its potential application. The majority of students believed that big data technology is practical in their clinical practice (77%), medical educations (85%), and scientific researches (82%). Over 80% of participants showed positive attitudes towards big data platforms. They also exhibited sufficient professionalism before the workshop. Meanwhile, the workshop has significantly promoted students' knowledge of medical big data ($P<0.05$), and lead to more positive attitudes to big data platforms and higher levels of professionalism.

Conclusions: Chinese medical students have primitive acquaintance and positive attitudes toward big data technologies. They also yield sufficient levels of professionalism. And the NUWA platform-based workshop was useful in furthering their understanding of big data and enhancing professionalism.

Keywords: Big data; Chinese medical education; Digital competencies; Professionalism; Workshop.

Strengths and limitations of this study

- This study depicted the cognition and professionalism of medical big data among Chinese medical students in the era of big data for the first time.
- This study provided a vivid example of how big data assists medical education, based on the first nation-wide gynecologic oncology medical big data platform in China, NUWA platform.
- The limited sample size may restrict the generalization of the conclusion.

Background

Electronification of medical records is the signature of the modern health care system.[1-3] Massive clinical and omics data were produced to enable more detailed depictions of patients and diseases. Medical big data is thus reshaping our appreciation of the modern medical system. Technological advancements in data storage, processing, and analysis accelerated the clinical application of big data-driven products and contributed to personalized disease management,[4, 5] early diagnosis,[6-8] and treatment decision.[9, 10] Especially in the COVID-19 pandemic setting, achieving rapid application for medical big data would meet the pressing clinical need to predict the progression of diseases by data characteristics.[11-13]

Meanwhile, big data has brought with its new challenges for doctors.[14] Information overload is posed to every healthcare worker, since they have to acclimate the nature of big data, including extraordinary value, volume, velocity, variety, and variability.[15] Challenges can also be expected on medical professionalism in the age of big data. Because big data valued realistic health-related information more than ever, the illusion that digital data overweigh face-to-face physician-patient interactions may motivate doctors to ignore the importance of professionalism.[16, 17] The situation may be worse in China, because of the explosive development in big data-based technology in the last years and the well-known heavy workloads of Chinese doctors.[18, 19] So extra lectures and workshops in obtaining insights into big data and remaining respectful for patients are necessary.

In this study, we carried on a workshop for Chinese clinical and pre-clinical students to further their knowledge of medical big data and to improve their professionalism in the big data era. The workshop was based on the National Union of Real-World Gynecologic Oncology Research & Patient Management Platform (NUWA platform), which is the first national-wide multi-function gynecologic oncology medical big data platform in China.

Methods

Study design

We performed a cross-sectional study on Chinese medical students' learning and application of big data in healthcare. The survey was conducted among students attending a workshop on medical big data.

Workshop design

The workshop was conducted on July 8, 2021 in Tongji Hospital, Tongji medical school, Huazhong University of Science and Technology, Wuhan, China. It is a part of the series class of "Medical big data platform learning and applying".

The workshop was delivered by two senior doctors (QLG and SQZ), who have over 5 years of experience in medical big data platform development and application. The workshop consisted of three major parts, (i) a lecture on the content and application of medical big data; (ii) a lecture on professionalism for doctors in the big data era; (iii) learners are allowed to explore the NUWA platform freely according to their interest. The first and second sections would last for about 45 minutes, and section three for 30 minutes.

In section one, five major parts were included in Dr. Zeng's speech, (i) the development of medical big data in China and the world, (ii) the application and potential of big data in clinical practice, medical education, and scientific researches; (iii) the construction and content of the NUWA platform, (iv) how to use the NUWA platform; (v) plans for the NUWA platform development. For the second section, Prof. Gao gave a lecture about (i) the attitude health care workers should have when communicating with patients and colleagues considering that big data have already reshaped our medical system, (ii) the importance of detailed and continuous patients' information for the development of big data platform, (iii) how to protect patient privacy on the databases and (iv) how to use NUWA platform to satisfy patients' needs. For section 3, all students would be given a temporary account for the NUWA platform and have the access to all deidentified patient information. They were allowed to

explore the medical data freely and view structured health care information for half an hour.

Data collection

Clinical and pre-clinical students in Tongji Medical school were invited to attend a 2-hour class by email or roadshows between June 8, 2021 and July 7, 2021. The contents and speakers for the workshop were presented in the email or during the roadshow. Besides, all 5th- (n=50, pre-clinical) and 6th-grade (n=49, clinical) students of 8-year undergraduate education in Tongji Medical School took part in this workshop as an additional course.

The participants should finish two surveys, one of which should be completed before and the other after the workshop. Both questionnaires contain their basic knowledge and attitude towards big data technology and a Professionalism Mini-Evaluation Exercise (P-MEX) instrument. The attitude towards big data platform was measured by 8 questions, namely, (i) Big data platform could assist future medical education; (ii) Big data platform could assist future medical researches; (iii)Big data platform could assist future clinical practice; (iv) I am willing to learn how to use big data platform; (v) I am willing to use big data platform in the future; (vi) I am willing to recommend big data platform to my colleagues; (vii) Big data platform could benefit my career; (viii) Big data platform could benefit all medical careers. For each question, students should choose from "Totally agree", "agree", "disagree", and "totally disagree". For the applicability of conclusions, "Totally agree" and "agree" were regarded as "positive attitude", "disagree" and "totally disagree" were considered as "negative attitude". The P-MEX consists of 24 items that represent four skill categories, including Doctor-Patient Relationship skills, Reflective skills, Time Management, and Inter-Professional Relationship skills.[20, 21] Besides, the first survey also contained students' baseline information, and the second one contained a satisfaction questionnaire. Meanwhile, the browsing histories for participants in section 3 were also recorded and analyzed to reflect students' interest.

The data collection and analysis plan were acknowledged and agreed by all participants at the beginning of the workshop and the study has been approved by the Research Ethics Commission of Tongji Hospital of Huazhong University of Science and Technology (2020-S201).

NUWA platform

The NUWA platform is the first nationwide Gynecological Oncology data-sharing platform launched by the National Clinical Research Center for Gynecological Oncology in August 2019. This platform integrated inpatient/outpatient clinical data, genomic data, and follow-up data to develop a patient-level longitudinal clinico-genomic database. Information was de-identified and extracted from electronic medical records. A rigorous data quality check was performed to ensure the accuracy of the data entries. Since its foundation in 2019, 17 first-class hospitals from different provinces or cities in China participated in the NUWA platform until August 2021.

Statistics analysis

Descriptive statistics were presented by counts and percentages to describe the demographic information. The Chi-square test was used to compare the changes in knowledge and understanding of big data and professionalism before and after the workshop. To acquire more practical results, the categories “not at all important” and “not important” were combined for analysis, as well as “important” and “very important”. *P-values <0.05* were considered significant. The data were analyzed using R Version 4.03.

Patient and Public Involvement

No patients or members of the public were involved in this study.

Results

Characteristics of students

A total of 274 students participated in this workshop and completed two surveys. All of them were included in the final analysis. Participants were aged between 22 and 28, with 148 (54%) pre-clinical students and 126 (46%) clinical ones. Among them, 130 (47.4%) were males and 144 (52.4%) were females (**Table 1**).

The majority of participants (207, 75.5%) used to know at least one big data platform in China or the world. And over two-thirds of them (183, 66.8%) also acknowledged its application. However, only 17.2% (47) of them have been involved in any project related to medical big data (**Table 1**).

Knowledge for big data platform

Before the workshop, about a quarter of the students (74, 27%) knew the detailed content of medical big data platforms while near 64% (174) knew the potential application of it. After attending the lessons, almost all students could understand the content (253, 92%) and the potential application (253, 92%) of medical big data (**Figure 1 and Table S1**).

Meanwhile, no matter before or after the workshop, the majority of students believed that big data technology is practical in medical educations, clinical practice, and scientific researches (85%, 77%, 82% before the workshop; 99%, 87%, 95% after the workshop, respectively) (**Figure 1 and Table S1**).

Students' attitude towards big data platform

Even before the workshop, most students hold a positive attitude on the potential of the big data platform in promoting medical educations (237, 86%), medical researches (256, 93%), and clinical practices (247, 90%). And they are also willing to learn and use the big data platform (250, 91%, and 218, 80% respectively). 74% (202) of them are enthusiastic to introduce big data platforms to their colleagues. Interestingly, most students were convinced that big data platform could benefit their careers (248, 91%), but they were not sure if it would yield the same effect on the others (128, 47%) (**Figure**

2 and Table S2).

When the workshop was finished, almost all students have positive attitudes towards the big data platform (**Figure 2 and Table S2**). However, there are still 12% (34) and 18% (48) of students who are not sure if the big data platform could benefit their and others' medical careers.

Professionalism

Generally, positive attitudes to all the professionalism items were demonstrated, with more than three-quarters of the students agree that all professionalism attributes are "important" or "very important" before the workshop. The three items with most students thought are "not at all important" or "not important" were maintaining patient confidentiality (66, 24% of students choose "not important" or "not at all important"), maintaining appropriate appearance (60, 22% students choose "not important" or "not at all important") and respecting rules and procedures of the system (55, 20% students choose "not important" or "not at all important") (**Figure 3A and Table S3**).

After the lectures, students exhibited a more positive attitude on most professionalism items ($p<0.05$). The most significant improvements happened on "Maintaining patient confidentiality" (from 76% to 95%), "Listen actively to patient" (from 89% to 100%), and "Accepting feedback" (from 89% to 98%) (**Figure 3B and Table S3**).

Interest in medical data

In the free exploration section, most students (253, 92.3%) viewed the patients' hospitalization logs. 76.3% (209) of them are interested in the history of illness. Meanwhile, many students were interested in medicine usages (56.2%), surgery reports (52.9%), and diseases of rare pathologies (58.6%). Only about one-third of students viewed the image diagnosis (37.6%) and follow-up records (32.5%) (**Table 2**).

Satisfactory survey

When finished, 95.6% (262) of participants were "satisfied" or "extremely satisfied" with this workshop, and only 7 out of 274 students were not at all satisfied. However, we failed to get feedback from them in the next three months. Most students thought the workshop informative (249, 90.9%) and understandable (255, 94.1%). The majority of students were also willing to recommend this seminar to other students (257, 93.8%) and to participant in similar classes in the future (254, 92.7%). As for the duration of this workshop, only 7 (2.6%) of students thought 2 hours was too long (**Table 3**).

Discussion

Our study demonstrated that Chinese medical students have primitive acquaintance and positive attitudes toward big data technologies. They also yield expertise in professionalism. Meanwhile, workshops based on big data platforms could further strengthen their digital competencies and improve doctor-patient communication capabilities, which would lead to better fitness during the expansion of medical big data.

In this workshop, the basic knowledge of big data platforms in China was measured. Although a large fraction of students knew about the big data platform and its utility, only a few of them could apprehend its composition or how to use it. This situation may be caused by the rapid development of big data technology in China recently and comparatively lagged relevant educations.[22, 23] Besides, ethical challenges that hinder medical and public health data sharing may also exacerbate the situation.[24] Therefore, aside from conducting big data-related workshops, promoting medical and public health data sharing policy and evolving relevant legal and ethical implications were also of great importance.

Students' enthusiasm towards medical big data was also taken into consideration. As expected and consistent with previous reports,[25] the majority of students exhibited positive attitudes towards the big data platform and were willing to acquire further proficiency. They believed that big data would play a vital role in future medical

education, clinical practice, and scientific researches, which are the main tasks for all Chinese doctors. However, not all of them believed big data could benefit medical careers in any case. The popularity of information technology in China recently may contribute to the big-data-friendly intention in medical students.[26] Meanwhile, their limited understanding of related fields may hamper the belief that it could be used productively.[27-29] In the meantime, we were delighted to see that almost all students recognized the significance of big data after our workshop. We thus believed that proper education may be an efficient way to diminish the misunderstanding and to achieve the full potential in big data platforms.

Professionalism was related to the development of big data in our study for the first time in China. Under the consideration that big data could reshape medical activities in all aspects including doctors' attitudes towards their colleagues and patients,[30] a lecture about "professionalism in the age of big data" was presented in the current workshop. Before the seminar, students demonstrated sufficient professionalism similar to previous reports, with high scores in almost all items. Several important professionalism levels were significantly increased in many elements after the workshop, indicating that this workshop could be a preliminary attempt for promoting professionalism when the digital data have changed the way we communicate and doctors have to spend more time with electronic records rather than patients. These changes our workshop brought deserve more attention in China since Chinese doctors were well-known to be overburdened.[18, 31] These improvements could also help increase the reliability of medical records and produce convincing and effective medical information.

Another interesting fact in the result is that about 60% of students noticed rare diseases in the free-exploration section, which is hardly involved in routine medical classes.[32, 33] We believe this is another strength of big data-based medical education. In traditional medical classes, it is arduous for teachers to grant detailed depictions for every type of rare disease. As a result, students may not have easy access to these exceptional cases, not to mention pay a visit to their medical records thoroughly. The

big data platform makes it possible for every user to browse cases with rare pathologies conveniently, which would benefit patients as well as reduce misdiagnosis.

Our workshop has some limitations. The most important one is that the NUWA platform is still under development and the omics data was not included currently. Therefore, another investigation should be performed when the construction of NUWA was completed. Furthermore, narrow geographical distributions of study participants interfere with the generalization of results to other populations. Meanwhile, there are concerns that students who volunteered to participate in the workshop may be more enthusiastic about big data technology than those who did not. We thus included all 5th- (pre-clinical) and 6th-grade (clinical) students of 8-year undergraduate education in Tongji Medical School to make the study sample more representative.

Conclusion

This study depicted Chinese students' knowledge of medical big data for the first time and the NUWA platform-based workshop was proven useful in promoting their understanding of big data and enhancing professionalism.

List of abbreviations

P-MEX, Professionalism Mini-Evaluation Exercise instrument; NUWA platform, the National Union of Real-World Gynecologic Oncology Research & Patient Management Platform.

Declarations

Ethics approval and consent to participate

The data collection and analysis plan were acknowledged and agreed by all participants and the study has been approved by the Research Ethics Commission of Tongji Hospital of Huazhong University of Science and Technology (2020-S201).

Consent for publication

Not applicable.

Availability of data and materials

The data used in the current study are available from the corresponding author on reasonable request via e-mail (qingleigao@hotmail.com).

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the Teaching Reform Project for Postgraduates of Huazhong University of Science and Technology (YY202006), and the Major Project of Medical Education (20A0769).

Authors' contributions

QLG, JHL and XFJ designed the study and survey. QLG, JHL, XFJ, ML, ZKP, YBH and SQZ conducted the workshop. PJ, HYL, JHC, XYL, YY, and SQZ analyzed and interpreted the data. YJZ and GCM performed interpretation of data and discussion of findings. QLG conceptualized and designed the study, supervised the project, analyzed and interpreted the data, and wrote the paper. All authors approved the final version of the manuscript and agreed to the submission of this manuscript.

Acknowledgements

Not applicable.

Reference

1. Bezemer T, de Groot MC, Blasse E, et al. A Human(e) Factor in Clinical Decision Support Systems. J Med Internet Res. 2019;21(3):e11732-e.
2. Sim I. Two Ways of Knowing: Big Data and Evidence-Based Medicine. Ann Intern Med. 2016;164(8):562-3.
3. Wu WT, Li YJ, Feng AZ, et al. Data mining in clinical big data: the frequently used databases, steps, and methodological models. Military Medical Research. 2021;8(1):44.
4. Olatosi B, Zhang J, Weissman S, et al. Using big data analytics to improve HIV medical care utilisation in South Carolina: A study protocol. #N/A. 2019;9(7):e027688.
5. Carra G, Salluh JIF, da Silva Ramos FJ, et al. Data-driven ICU management: Using Big Data and algorithms to improve outcomes. J Crit Care. 2020;60:300-4.
6. Kumar NMS, Eswari T, Sampath P, et al. Predictive Methodology for Diabetic Data Analysis in Big Data. In: Vijayakumar V, Neelanarayanan V, editors. Big Data, Cloud and Computing Challenges. Procedia Computer Science. 502015. p. 203-8.

7. Razavian N, Blecker S, Schmidt AM, et al. Population-Level Prediction of Type 2 Diabetes From Claims Data and Analysis of Risk Factors. #N/A. 2015;3(4):277-87.
8. Fitzgerald RC. Big data is crucial to the early detection of cancer. Nat Med. 2020;26(1):19-20.
9. Fridley BL, Koestler DC, Godwin AK. Individualizing care for ovarian cancer patients using big data. J Natl Cancer Inst. 2014;106(5).
10. Cammarota G, Ianiro G, Ahern A, et al. Gut microbiome, big data and machine learning to promote precision medicine for cancer. Nat Rev Gastroenterol Hepatol. 2020;17(10):635-48.
11. Harmon SA, Sanford TH, Xu S, et al. Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets. Nat Commun. 2020;11(1):4080.
12. Mei X, Lee HC, Diao KY, et al. Artificial intelligence-enabled rapid diagnosis of patients with COVID-19. Nat Med. 2020;26(8):1224-8.
13. Wang CJ, Ng CY, Brook RH. Response to COVID-19 in Taiwan: Big Data Analytics, New Technology, and Proactive Testing. Jama. 2020;323(14):1341-2.
14. Jensen PB, Jensen LJ, Brunak S. Mining electronic health records: towards better research applications and clinical care. Nature reviews Genetics. 2012;13(6):395-405.
15. Andreu-Perez J, Poon CC, Merrifield RD, et al. Big data for health. IEEE J Biomed Health Inform. 2015;19(4):1193-208.
16. O'Connor S, Zhang M, Honey M, et al. Digital professionalism on social media: A narrative review of the medical, nursing, and allied health education literature. International journal of medical informatics. 2021;153:104514.
17. Attai DJ, Anderson PF, Fisch MJ, et al. Risks and benefits of Twitter use by hematologists/oncologists in the era of digital medicine. Seminars in hematology. 2017;54(4):198-204.
18. Jingang A. Which future for doctors in China? Lancet. 2013;382(9896):936-7.
19. Wu D, Wang Y, Lam KF, et al. Health system reforms, violence against doctors and job satisfaction in the medical profession: a cross-sectional survey in Zhejiang Province, Eastern China. BMJ open. 2014;4(12):e006431-e.
20. Cruess R, McIlroy JH, Cruess S, et al. The Professionalism Mini-Evaluation Exercise: A Preliminary Investigation. Academic Medicine. 2006;81(10).
21. Norcini JJ, Blank LL, Arnold GK, et al. The mini-CEX (clinical evaluation exercise): a preliminary investigation. Ann Intern Med. 1995;123(10):795-9.
22. Wu J, Wang J, Nicholas S, et al. Application of Big Data Technology for COVID-19 Prevention and Control in China: Lessons and Recommendations. J Med Internet Res. 2020;22(10):e21980.
23. Shen T, Wang C. Big Data Technology Applications and the Right to Health in China during the COVID-19 Pandemic. International journal of environmental research and public health. 2021;18(14).
24. Li X, Cong Y. A Systematic Literature Review of Ethical Challenges Related to Medical and Public Health Data Sharing in China. Journal of Empirical Research on Human Research Ethics. 2021:15562646211040299.
25. Parthiban N, Boland F, Fadil Azim DH, et al. Asian medical students' attitudes towards professionalism. Med Educ Online. 2021;26(1):1927466.
26. Liang W, Yao J, Chen A, et al. Early triage of critically ill COVID-19 patients using deep learning. Nature communications. 2020;11(1):3543-.
27. Meo SA, Sattar K, Ullah CH, et al. Progress and prospects of medical education research in Asian Countries. Pakistan journal of medical sciences. 2019;35(6):1475-81.
28. Schmid RD, Xiong X. Biotech in China 2021, at the beginning of the 14th five-year period ("145").

Applied microbiology and biotechnology. 2021;105(10):3971-85.
29. Wu H, Li W, Xie A, et al. Funding of health professional education: China's 20-year process and a global comparison. Medical education. 2021.
30. Silva Junior G, Askari M, Oliveira J. Digital Health and Possible Solutions to Improve the Care in the Field of Nephrology. Contributions to nephrology. 2021;199:1-15.
31. Liu S, Chen Y, Li S, et al. What Are the Important Factors Influencing the Recruitment and Retention of Doctoral Students in a Public Health Setting? A Discrete Choice Experiment Survey in China. International journal of environmental research and public health. 2021;18(18).
32. Walkowiak D, Domaradzki J. Are rare diseases overlooked by medical education? Awareness of rare diseases among physicians in Poland: an explanatory study. Orphanet journal of rare diseases. 2021;16(1):400.
33. Domaradzki J, Walkowiak D. Knowledge and Attitudes of Future Healthcare Professionals Toward Rare Diseases. Front Genet. 2021;12:639610.

Tables

Table 1. Baseline characteristics for participants.

Characteristics	Number of participants
Age, years	
22 (n, %)	27, 9.9%
23 (n, %)	16, 5.8%
24 (n, %)	67, 24.5%
25 (n, %)	66, 24.1%
26 (n, %)	50, 18.2%
27 (n, %)	28, 10.2%
28 (n, %)	20, 7.3%
Gender	
male (n, %)	130, 47.4%
female (n, %)	144, 52.6%
Study stage	
pre-clinical (n, %)	148, 54%
clinical (n, %)	126, 46%
Acknowledgment of any kind of big data platform	
yes (n, %)	207, 75.5%
no (n, %)	67, 24.5%
Know the applications of big data technology	
yes (n, %)	183, 66.8%
no (n, %)	91, 33.2%
Involved in any big data-related projects	
yes (n, %)	47, 17.2%
no (n, %)	227, 82.8%
Total	274, 100%

Table 2. Browser records in the free-exploration section.

Content	Number of participants
History of Illness	
Yes (n, %)	209, 76.3%
no (n, %)	65, 23.7%
Hospitalization logs	
yes (n, %)	253, 92.3%
no (n, %)	21, 7.7%
Medicine usage	
yes (n, %)	154, 56.2%
no (n, %)	120, 43.8%
Surgery record	
yes (n, %)	145, 52.9%
no (n, %)	129, 47.1%
Image diagnosis	
yes (n, %)	103, 37.6%
no (n, %)	171, 62.4%
Follow-up records	
yes (n, %)	89, 32.5%
no (n, %)	185, 67.5%
Rare pathologies	
yes (n, %)	160, 58.4%
no (n, %)	114, 41.6%
Total (n, %)	274, 100%

Table 3. Participants 'answers to the workshop satisfaction survey

	Not at all (n, %)	No (n, %)	Yes (n, %)	Yes, extremely (n, %)
Overall, are you satisfied with this course	7, 2.6%	5, 1.8%	57, 20.8%	205, 74.8%
Did you think the course are informative?	10, 3.6%	15, 5.5%	101, 36.9%	148, 54%
Did you think the duration of this courses is too long	121, 44.2%	118, 43.1%	28, 10.2%	7, 2.6%
Was the course understandable for you	11, 4%	8, 2.9%	121, 44.2%	134, 48.9%
Would you recommend these courses to other students?	6, 2.2%	11, 4%	144, 52.6%	113, 41.2%
Are you willing to take part in similar courses in the future?	6, 2.2%	14, 5.1%	35, 12.8%	219, 79.9%

Figure Legends

Figure 1. Basic knowledge for big data platform (A) before and (B) after the workshop.

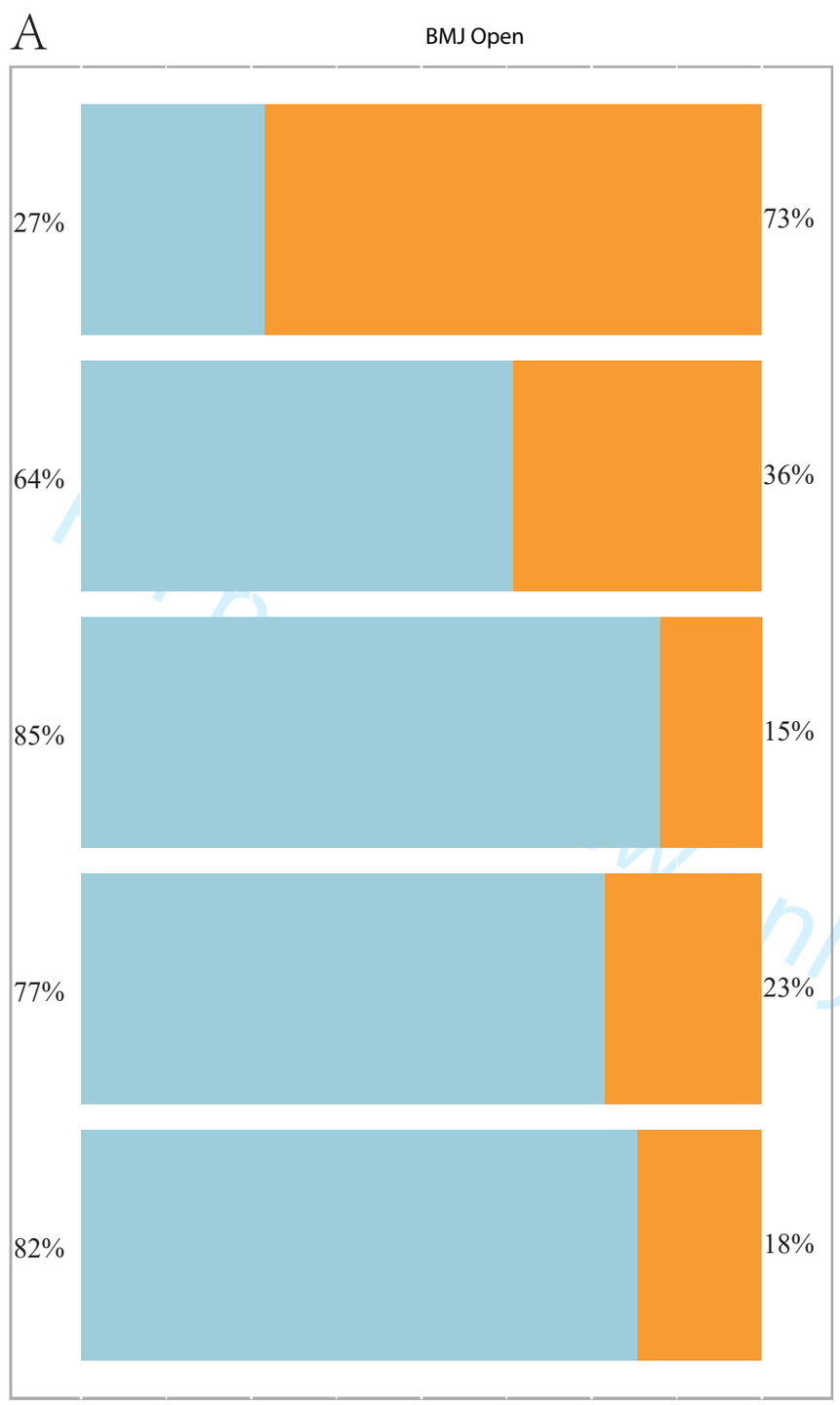
* means that there is a significant difference before and after the workshop.

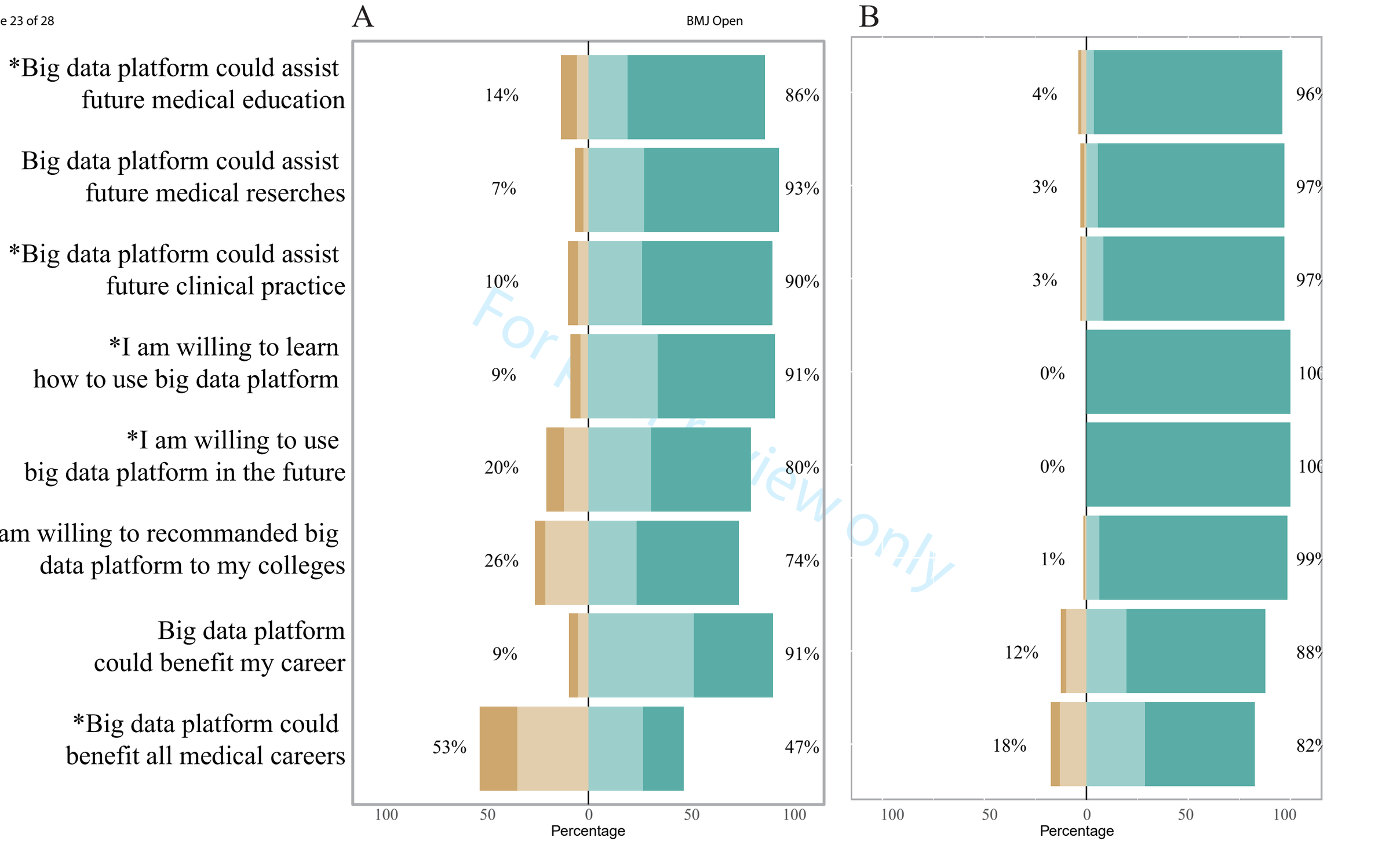
Figure 2. Students' attitude towards big data platform (A) before and (B) after the workshop.

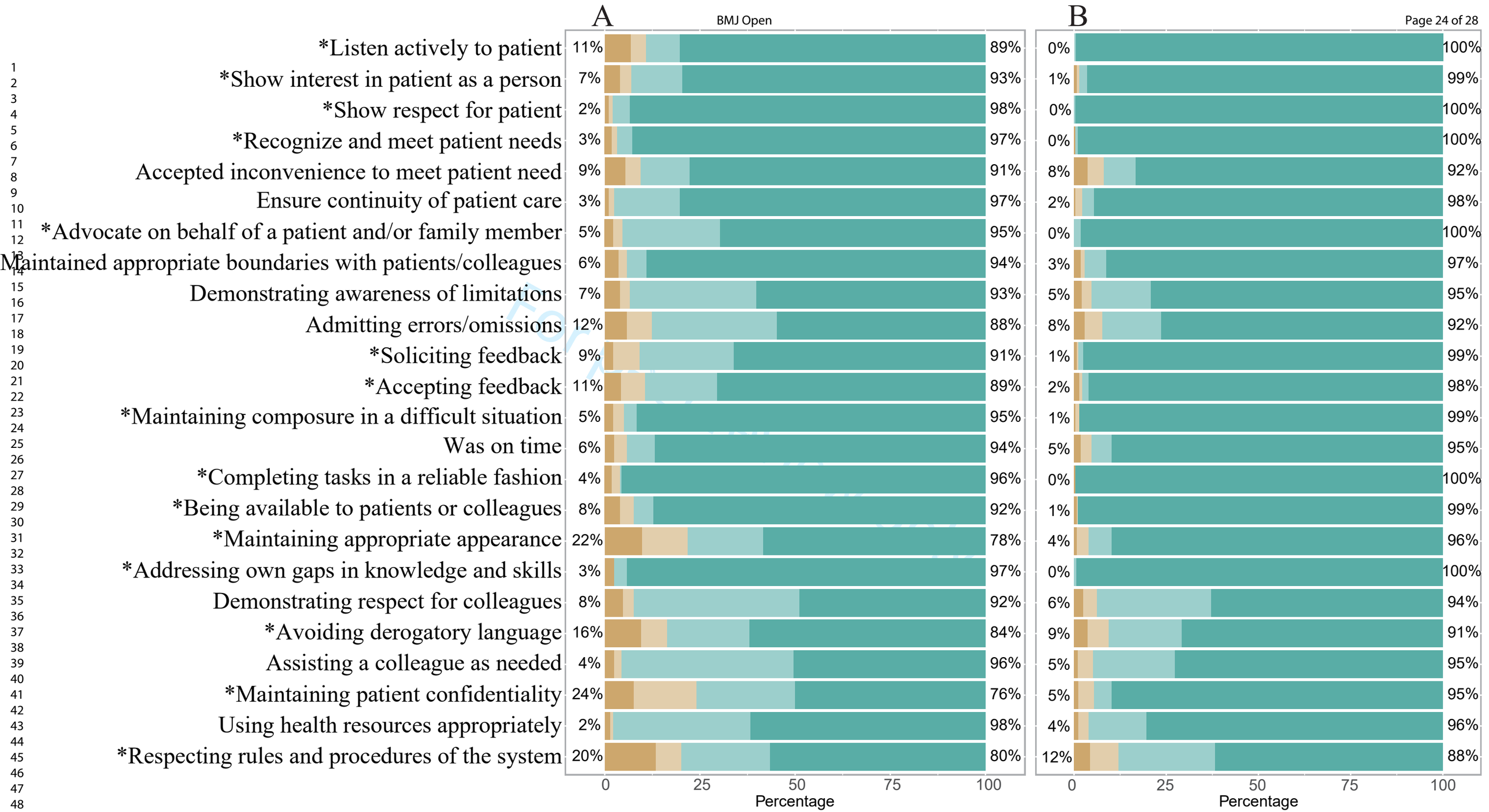
* means that there is a significant difference before and after the workshop.

Figure 3. The professionalism for students (A) before and (B) after the workshop.

* means that there is a significant difference before and after the workshop.







Supplemental tables

Table S1, basic knowledge for big data platform before and after the workshop.

	Before Workshop		After Workshop		
Items	YES, n	NO, n	YES, n	NO, n	*P*
I know the content of big data platform	74	200	253	21	*<0.0001*
I know the potential applications of big data technology	174	100	253	21	*<0.0001*
Big data technology is practical in medical education	233	41	271	3	*<0.0001*
Big data technology is practical in clinical practice	211	63	238	36	*0.0039*
Big data technology is practical in scientific researches	224	50	261	13	*<0.0001*

Table S2, students' attitude towards big data platform before and after the workshop.

	Before Workshop				After Workshop				
Items	Totally disagree, n	Disagree, n	Agree, n	Totally agree, n	Totally disagree, n	Disagree, n	Agree, n	Totally agree, n	*P*
Big data platform could assist future medical education	21	16	53	184	4	7	10	253	*0.0002*
Big data platform could assist future scientific researches	11	7	75	181	5	3	16	250	*0.0705*
Big data platform could assist future clinical practice	13	14	72	175	2	6	23	243	*0.0017*
I am willing to learn how to use big data platform	13	11	93	157	0	0	0	274	*<0.0001*
I am willing to use big data platform in the future	23	33	84	134	0	0	0	274	*<0.0001*
I am willing to recommended big data platform to my colleagues	14	58	65	137	2	2	18	252	*<0.0001*
Big data platform could benefit my career	12	14	142	106	7	27	54	186	*0.3382*
Big data platform could benefit all medical careers	50	96	74	54	12	36	79	147	*<0.0001*

Table S3, the professionalism for students before and after the workshop.

	Before workshop				After workshop				
Items	Not at all important, n	Not important, n	Important, n	Very important, n	Not at all important, n	Not important, n	Important, n	Very important, n	*P*
Doctor-patient relationship skills									
Listen actively to patient	19	11	24	220	0	0	1	273	*<0.000 1*
Show interest in patient as a person	11	8	37	218	2	2	6	264	*0.0029*
Show respect for patient	3	3	12	256	0	0	1	273	*0.0401*
Recognize and meet patient needs	5	4	11	254	1	0	2	271	*0.0255*
Accepted inconvenience to meet patient need	15	11	35	213	10	12	24	228	*0.6503*
Ensure continuity of patient care	3	4	47	220	1	5	9	259	*1*
Advocate on behalf of a patient and/or family member	6	7	70	191	0	0	5	269	*0.0008*
Maintained appropriate boundaries with patients/colleagues	10	6	14	244	5	3	16	250	*0.1440*
Reflective skills									
Demonstrating awareness of limitations	11	7	91	165	6	7	44	217	*0.4595*
Admitting errors/omissions	16	18	90	150	8	13	44	209	*0.0880*
Soliciting feedback	6	19	68	181	2	1	4	267	*<0.000*

									1
Accepting feedback	12	17	52	193	4	2	5	263	*0.0001*
Maintaining composure in a difficult situation	6	8	9	251	1	3	0	270	*0.0310*
Time management									
Was on time	7	9	20	238	5	8	15	246	*0.7027*
Completing tasks in a reliable fashion	5	6	1	262	1	0	0	273	*0.0086*
Being available to patients or colleagues	11	10	14	239	2	1	0	271	*0.0004*
Interprofessional relationship skills									
Maintaining appropriate appearance	27	33	54	160	2	9	17	246	*<0.000 1*
Addressing own gaps in knowledge and skills	7	0	9	258	0	0	2	272	*0.0225*
Demonstrating respect for colleagues	13	8	119	134	7	10	85	172	*0.6139*
Avoiding derogatory language	26	19	59	170	10	16	54	194	*0.0220*
Assisting a colleague as needed	7	5	124	138	3	11	61	199	*0.8407*
Maintaining patient confidentiality	21	45	71	137	3	12	13	246	*<0.000 1*
Using health resources appropriately	4	2	99	169	3	8	43	220	*0.3244*
Respecting rules and procedures of the system	37	18	64	155	12	21	72	169	*0.0146*

STROBE 2007 (v4) Statement—Checklist of items that should be included in reports of *cross-sectional studies*

Section/Topic	Item #	Recommendation	Reported on page #
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4
Objectives	3	State specific objectives, including any prespecified hypotheses	4
Methods			
Study design	4	Present key elements of study design early in the paper	5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5-6
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants	5
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	6
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	6
Bias	9	Describe any efforts to address potential sources of bias	NA
Study size	10	Explain how the study size was arrived at	6
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	6
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	7
		(*b*) Describe any methods used to examine subgroups and interactions	NA
		(*c*) Explain how missing data were addressed	7
		(*d*) If applicable, describe analytical methods taking account of sampling strategy	7
		(*e*) Describe any sensitivity analyses	NA
Results			

Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	8
		(b) Give reasons for non-participation at each stage	All participants were eligible and included in this study
		(c) Consider use of a flow diagram	NA
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	8
		(b) Indicate number of participants with missing data for each variable of interest	Very high response rates with no missing data, as shown in supplementary data.
Outcome data	15*	Report numbers of outcome events or summary measures	8
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	In all tables and figures.
		(*b*) Report category boundaries when continuous variables were categorized	In all tables and figures.
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	NA
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	NA
Discussion			
Key results	18	Summarise key results with reference to study objectives	10
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	12
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	12
Generalisability	21	Discuss the generalisability (external validity) of the study results	12
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	13

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

Oncologic big data platform for promoting digital competencies and professionalism in Chinese medical students: a longitudinal study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2022-061015.R1
Article Type:	Original research
Date Submitted by the Author:	13-Jul-2022
Complete List of Authors:	Liu, Jiahao; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Jiao, Xiaofei; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Zeng, Shaoqing; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Li, Huayi; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Jin, Ping; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Chi, Jianhua; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Liu, Xingyu; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Yu, Yang; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Ma, Guanchen; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Zhao, Yingjun; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Li, Ming; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Peng, Zikun; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Huo, Yabing; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Gao, Qing-Lei ; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology
Primary Subject Heading:	Medical education and training
Secondary Subject Heading:	Medical education and training
Keywords:	MEDICAL EDUCATION & TRAINING, Gynaecological oncology < GYNAECOLOGY, Gynaecological oncology < ONCOLOGY, Health informatics < BIOTECHNOLOGY & BIOINFORMATICS

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Oncologic big data platform for promoting digital competencies and professionalism in Chinese medical students: a longitudinal study

Jiahao Liu[1,2#], Xiaofei Jiao[1,2#], Shaoqing Zeng[1,2], Huayi Li[1,2], Ping Jin[1,2], Jianhua Chi[1,2], Xingyu Liu[1,2], Yang Yu[1,2], Guanchen Ma[1,2], Yingjun Zhao[1,2], Ming Li[1,2], Zikun Peng[1,2], Yabing Huo[1,2], Qinglei Gao[1,2*]

[1] Department of Obstetrics and Gynecology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, Hubei, P. R. China.
[2] Cancer Biology Research Center (Key Laboratory of the Ministry of Education), Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, Hubei, P. R. China.

[#] These authors contribute equally.

* Correspondence to:
Prof. Qinglei Gao, Cancer Biology Research Center (Key Laboratory of the Ministry of Education), Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, 1095 Jiefang Ave, Wuhan 430030, Hubei, P. R. China. Tel: +86-27-83662681; Fax: +86-27-83662681; E-mail: qingleigao@hotmail.com.

Word count (excluding title page, abstract, references, figures and tables): 3,162

Abstract

Objectives: Advancements in big data technology are reshaping the health care system in China. This study aims to explore the role of medical big data in promoting digital competencies and professionalism among Chinese medical students.

Design, setting and participants: This study was conducted among 274 medical students who attended a workshop on medical big data conducted on July 8, 2021 in Tongji Hospital. The workshop was based on the first nationwide multifunction gynecologic oncology medical big data platform in China, at the National Union of Real-World Gynecologic Oncology Research & Patient Management Platform (NUWA platform).

Outcome measures: Data on knowledge, attitudes toward big data technology and professionalism were collected before and after the workshop. We have measured the four skill categories: doctor–patient relationship skills, reflective skills, time management, and interprofessional relationship skills using the Professionalism Mini–Evaluation Exercise (P-MEX) as a reflection for professionalism.

Results: A total of 274 students participated in this workshop and completed all the surveys. Before the workshop, only 27% of them knew the detailed content of medical big data platforms, and 64% knew the potential application of medical big data. The majority of the students believed that big data technology is practical in their clinical practice (77%), medical education (85%), and scientific research (82%). Over 80% of the participants showed positive attitudes toward big data platforms. They also exhibited sufficient professionalism before the workshop. Meanwhile, the workshop significantly promoted students' knowledge of medical big data ($P < 0.05$), and led to more positive attitudes toward big data platforms and higher levels of professionalism.

Conclusions: Chinese medical students have primitive acquaintance and positive attitudes toward big data technologies. The NUWA platform-based workshop may potentially promote their understanding of big data and enhance professionalism, according to the self-measured P-MEX scale.

Keywords: Big data, Chinese medical education, Digital competencies,

Professionalism, Workshop

Strengths and limitations of this study

- This study depicted the knowledge and professionalism of medical big data among Chinese medical students in the era of big data for the first time.
- This study provided a vivid example of how big data can assist medical education, based on the first nationwide gynecologic oncology medical big data platform in China (NUWA platform).
- The generalization of the findings is limited due to the sample representation and longitudinal continuity of this study.

Background

Electronification of medical records is the signature of the modern health care system.[1-3] Massive clinical and omics data have been produced to enable more detailed depictions of patients and diseases. Medical big data is thus reshaping our appreciation of the modern medical system. Technological advancements in data storage, processing, and analysis accelerated the clinical application of big data-driven products and contributed to personalized disease management,[4, 5] early diagnosis,[6-8] and treatment decision.[9, 10] Especially in the COVID-19 pandemic setting, achieving rapid application of medical big data would meet the pressing clinical need to predict the progression of diseases using data characteristics.[11-13]

Big data has brought new challenges for doctors.[14] Information overload is a challenge for every healthcare worker, since they have to acclimate to the nature of big data, including extraordinary value, volume, velocity, variety, and variability.[15] Challenges can also be expected regarding medical professionalism in the age of big data. Because big data values realistic health-related information more than ever, the illusion that digital data outweigh face-to-face physician–patient interactions may motivate doctors to ignore the importance of professionalism.[16, 17] There is no precise definition of medical professionalism, which is reflected in the attitudes and behaviors directly related to clinical practice. Epstein RM et al. proposed the definition of professional competence as wisely using communication, knowledge, technical skills, clinical reasoning, emotions, values, and reflection in daily practice.[18] And Cain J et al. summarized professionalism as a series of attributes, such as altruism, respect, honesty, and so on.[19] Some studies have shown that digital medical education based on big data plays a positive effect in promoting professionalism.[16, 20, 21] However, there is a lack of research on the attitude and professionalism of Chinese healthcare workers regarding medical big data. Therefore, extra lectures and workshops about obtaining insights into big data and remaining respectful to patients are necessary.

In this study, we aimed at depicting the knowledge of medical big data and professionalism in the era of big data among Chinese medical students. What's more,

we conducted a workshop for Chinese clinical and preclinical students to improve their knowledge of medical big data and their professionalism in the big data era, based on the National Union of Real-World Gynecologic Oncology Research & Patient Management Platform (NUWA platform), which is the first nationwide multifunction gynecologic oncology medical big data platform in China.

Methods

Study design

A workshop on the introduction of medical big data was conducted among clinical and preclinical students in Tongji Medical School. We conducted this study on Chinese medical students' learning and the application of big data in healthcare using pre- and post-course questionnaires.

Workshop design

The workshop was conducted on July 8, 2021 in Tongji Hospital, Tongji Medical School, Huazhong University of Science and Technology, Wuhan, China. It was part of the series class of "Medical big data platform learning and applying".

The workshop was delivered by two senior doctors (QLG and SQZ), who have over 5 years of experience in medical big data platform development and application. The workshop consisted of three major parts, (i) a lecture on the content and application of medical big data; (ii) a lecture on professionalism for doctors in the big data era; and (iii) learners being allowed to explore the NUWA platform freely according to their interest. The first and second sections lasted for approximately 45 minutes, and the third section lasted for 30 minutes.

In section one, five major parts were included in lectures given by the senior doctors, (i) the development of medical big data in China and the world, (ii) the application and potential of big data in clinical practice, medical education, and scientific research, (iii) the construction and content of the NUWA platform, (iv) how to use the NUWA

platform, and (v) plans for the NUWA platform development. For the second section, the following four items were discussed: (i) the attitude health care workers should have when communicating with patients and colleagues considering that big data have already reshaped our medical system, (ii) the importance of detailed and continuous patient information for the development of big data platforms, (iii) how to protect patient privacy on the databases, and (iv) how to use the NUWA platform to satisfy patients' needs. For Section 3, all students were given a temporary account for the NUWA platform and have access to all deidentified patient information. They were allowed to explore the medical data freely and view structured health care information for half an hour.

Data collection

Clinical and preclinical students in Tongji Medical School were invited to attend a 2-hour class by email or roadshows between June 8, 2021, and July 7, 2021. The contents and speakers for the workshop were presented in the email or during the roadshow. In addition, all 5th- (n=50, preclinical) and 6th-grade (n=49, clinical) students of 8-year undergraduate education in Tongji Medical School took part in this workshop as an additional course.

The participants completed two surveys, one before the workshop and the other after the workshop. Both questionnaires contain their basic knowledge and attitude toward big data technology and a Professionalism Mini-Evaluation Exercise (P-MEX) instrument. The P-MEX is developed from the mini-Clinical Examination Exercise (mini-CEX) format by Cruess R et al. in 2006 to evaluate professionalism in clinical training.[22] It consists of 24 items representing four skill categories: doctor–patient relationship skills, reflective skills, time management, and interprofessional relationship skills skills.[23, 24] And the reliability and validity of P-MEX have been confirmed in both Eastern and Western cultural backgrounds.[25-27] The attitude scale is a self-created scale focused on measuring participants' pedagogic evaluation, acceptance and expectation of the big data platform. It consisted of 8 questions: (i) Big

data platform could assist future medical education, (ii) Big data platform could assist future medical research, (iii)Big data platform could assist future clinical practice, (iv) I am willing to learn how to use big data platform, (v) I am willing to use big data platform in the future, (vi) I am willing to recommend big data platform to my colleagues, (vii) Big data platform could benefit my career, and (viii) Big data platform could benefit all medical careers. For each question, students chose from "totally agree", "agree", "disagree", and "totally disagree". For the applicability of conclusions, "totally agree" and "agree" were regarded as "positive attitude", and "disagree" and "totally disagree" were considered as "negative attitude". In addition, the first survey also contained students' baseline information, and the second survey contained a satisfaction questionnaire. Meanwhile, the browsing histories for participants in Section 3 were also recorded and analyzed to reflect students' interest.

The data collection and analysis plan were acknowledged and agreed upon by all participants at the beginning of the workshop, and the study was approved by the Research Ethics Commission of Tongji Hospital of Huazhong University of Science and Technology (2020-S201).

NUWA platform

The NUWA platform is the first nationwide Gynecological Oncology data-sharing platform launched by the National Clinical Research Center for Gynecological Oncology in August 2019. This platform integrated inpatient/outpatient clinical data, genomic data, and follow-up data to develop a patient-level longitudinal clinico-genomic database. Information was deidentified and extracted from electronic medical records. A rigorous data quality check was performed to ensure the accuracy of the data entries. Since its foundation in 2019, 17 first-class hospitals from different provinces or cities in China have participated in the NUWA platform until August 2021.

Statistical analysis

Descriptive statistics were presented by counts and percentages to describe the

demographic information. The Chi-square test was used to compare the changes in knowledge and understanding of big data and professionalism before and after the workshop. To acquire more practical results, the categories "not at all important" and "not important" were combined for the analysis, as well as "important" and "very important". *P values <0.05* were considered statistically significant. The data were analyzed using R Version 4.03.

Patient and Public Involvement

No patients or members of the public were involved in this study.

Results

Characteristics of students

A total of 274 students participated in this workshop and completed two surveys. All of them were included in the final analysis. Participants were aged between 22 and 28, with 148 (54.0%) preclinical students and 126 (46.0%) clinical students. Among them, 130 (47.4%) were males and 144 (52.4%) were females (**Table 1**).

The majority of participants (207, 75.5%) knew of at least one big data platform in China or the world. Over two-thirds of them (183, 66.8%) also acknowledged its application. However, only 17.2% (47) of them had been involved in any project related to medical big data (**Table 1**).

Table 1. Baseline characteristics for participants.

Characteristics	Number of participants
Age, years	
22 (n, %)	27, 9.9%
23 (n, %)	16, 5.8%
24 (n, %)	67, 24.5%
25 (n, %)	66, 24.1%
26 (n, %)	50, 18.2%

	27 (n, %)	28, 10.2%
	28 (n, %)	20, 7.3%
Gender		
	male (n, %)	130, 47.4%
	female (n, %)	144, 52.6%
Study stage		
	pre-clinical (n, %)	148, 54.0%
	clinical (n, %)	126, 46.0%
Acknowledgment of any kind of big data platform		
	yes (n, %)	207, 75.5%
	no (n, %)	67, 24.5%
Know the applications of big data technology		
	yes (n, %)	183, 66.8%
	no (n, %)	91, 33.2%
Involved in any big data-related projects		
	yes (n, %)	47, 17.2%
	no (n, %)	227, 82.8%
Total		274, 100%

Knowledge of big data platforms

Before the workshop, approximately a quarter of the students (74, 27%) knew detailed content of at least one medical big data platform, while nearly 64% (174) knew the potential application of medical big data platforms. After attending the lessons, almost all students could understand the content (253, 92%) and the potential application (253, 92%) of medical big data (**Figure 1 and Table S1**).

Regardless of the survey completed (pre- or post-course), the majority of students believed that big data technology is practical in medical education, clinical practice,

and scientific research (85%, 77%, 82% before the workshop; 99%, 87%, 95% after the workshop, respectively) (**Figure 1 and Table S1**).

Students' attitudes toward big data platforms

Even before the workshop, most students hold a positive attitude on the potential of the big data platform in promoting medical education (237, 86%), medical research (256, 93%), and clinical practices (247, 90%). They were also willing to learn about and use the big data platform (250, 91%, and 218, 80% respectively). Of them, 74% (202) were enthusiastic to introduce big data platforms to their colleagues. Interestingly, most students were convinced that the big data platform could benefit their careers (248, 91%), but they were not sure if it would yield the same effect on the others (128, 47%) (**Figure 2 and Table S2**).

When the workshop was completed, almost all students had positive attitudes toward the big data platform (**Figure 2 and Table S2**). However, there were still 12% (34) and 18% (48) of students who were not sure if the big data platform could benefit their and others' medical careers, respectively.

Professionalism

Generally, positive attitudes toward all the professionalism items were demonstrated, with more than three-quarters of the students agreeing that all professionalism attributes were "important" or "very important" before the workshop. The three items that most students thought were "not at all important" or "not important" were maintaining patient confidentiality (66, 24% of students chose "not important" or "not at all important"), maintaining appropriate appearance (60, 22% students chose "not important" or "not at all important") and respecting rules and procedures of the system (55, 20% students chose "not important" or "not at all important") (**Figure 3A and Table S3**).

After the lectures, students exhibited a more positive attitude on most professionalism items ($p<0.05$). The most significant improvements occurred for

"Maintaining patient confidentiality" (from 76% to 95%), "Listen actively to patient" (from 89% to 100%), and "Accepting feedback" (from 89% to 98%) (**Figure 3B and Table S3**).

Interest in medical data

In the free exploration section, most students (253, 92.3%) viewed the patients' hospitalization logs. A total of 76.3% (209) of students were interested in the patients' history of illness. Meanwhile, many students were interested in medicine usages (56.2%), surgery reports (52.9%), and diseases of rare pathologies (58.6%). Only approximately one-third of students viewed image diagnosis (37.6%) and follow-up records (32.5%) (**Table 2**).

Table 2. Browser records in the free-exploration section.

Content	Number of participants
History of Illness	
Yes (n, %)	209, 76.3%
no (n, %)	65, 23.7%
Hospitalization logs	
yes (n, %)	253, 92.3%
no (n, %)	21, 7.7%
Medicine usage	
yes (n, %)	154, 56.2%
no (n, %)	120, 43.8%
Surgery record	
yes (n, %)	145, 52.9%
no (n, %)	129, 47.1%
Image diagnosis	
yes (n, %)	103, 37.6%
no (n, %)	171, 62.4%

Follow-up records	
yes (n, %)	89, 32.5%
no (n, %)	185, 67.5%
Rare pathologies	
yes (n, %)	160, 58.4%
no (n, %)	114, 41.6%
Total (n, %)	274, 100%

Satisfactory survey

Upon completing the survey, 95.6% (262) of participants were "satisfied" or "extremely satisfied" with this workshop, and only 7 out of 274 students were not at all satisfied. However, we failed to obtain feedback from them in the next three months. Most students thought that the workshop was informative (249, 90.9%) and understandable (255, 94.1%). The majority of students were also willing to recommend this seminar to other students (257, 93.8%) and to participate in similar classes in the future (254, 92.7%). Regarding the duration of this workshop, only 7 (2.6%) students thought that 2 hours was too long (**Table 3**).

Table 3. Participants 'answers to the workshop satisfaction survey

	Not at all (n, %)	No (n, %)	Yes (n, %)	Yes, extremely (n, %)	Mean	SD
Overall, are you satisfied with this course	7, 2.6%	5, 1.8%	57, 20.8%	205, 74.8%	3.68	0.640
Did you think the course are informative?	10, 3.6%	15, 5.5%	101, 36.9%	148, 54%	3.41	0.757
Did you think the duration of this courses is too long	121, 44.2%	118, 43.1%	28, 10.2%	7, 2.6%	2.07	0.954
Was the course	11, 4%	8, 2.9%	121 44.2%	134, 48.9%	3.38	0.733

understandable for you			,			
Would you recommend these courses to other students?	6, 2.2%	11, 4%	144 , 52.6%	113, 41.2%	3.33	0.659
Are you willing to take part in similar courses in the future?	6, 2.2%	14, 5.1%	35, 12.8%	219, 79.9%	3.70	0.666

SD, Standard Deviation.

Discussion

Our study demonstrated that Chinese medical students have little knowledge of and positive attitudes toward big data technologies. They also yield expertise in professionalism. Furthermore, workshops based on big data platforms could further strengthen their digital competencies and improve doctor–patient communication capabilities, which would lead to better fitness during the expansion of medical big data.

In this workshop, the basic knowledge of big data platforms in China was measured. Although a large fraction of students knew about the big data platform and its utility, only a few of them could apprehend its composition or how to use it. This situation may be caused by the recent rapid development of big data technology in China and comparatively lagging relevant education.[28, 29] In addition, ethical challenges that hinder medical and public health data sharing may have also exacerbated the situation.[30] Therefore, aside from conducting big data-related workshops, promoting medical and public health data sharing policies and evolving relevant legal and ethical implications are also of great importance.

Students' enthusiasm toward medical big data was also taken into consideration. As expected and consistent with previous reports,[25] the majority of students exhibited positive attitudes toward the big data platform and were willing to acquire further proficiency. They believed that big data would play a vital role in future medical

education, clinical practice, and scientific research, which are the main tasks for all Chinese doctors. However, not all of them believed that big data could benefit all medical careers. The popularity of information technology of China recently may contribute to the big data-friendly intention in medical students.[31] Meanwhile, their limited understanding of related fields may hamper the belief that it could be used productively.[32-34] In the meantime, we were delighted to see that almost all students recognized the significance of big data after our workshop. We thus believe that proper education may be an efficient way to diminish misunderstandings and to achieve full potential in big data platforms.

Professionalism was related to the development of big data in our study for the first time in China. Considering that big data can reshape medical activities in all aspects, including doctors' attitudes toward their colleagues and patients,[35] a lecture about "professionalism in the age of big data" was presented in the current workshop. Before the seminar, students demonstrated sufficient professionalism similar to the findings of previous reports, with high scores on the majority of these items. Nevertheless, it is worth noting that nearly 1/4 of the participants did not pay enough attention to maintaining patient confidentiality. Protecting patients' privacy is an essential embodiment of medical ethics and humanities.[36] Participants mostly answered these questions from the perspective of big data users before our workshop, ignoring that those data represented thousands of actual patients. After we emphasized the importance of data privacy in the workshop, students realized that respect for patients is the foundation for improving their medical professionalism. Meanwhile, to enhance patient-data privacy protection, the privacy information, such as ID number, was privatized and converted into an alternative ID number using a hashing algorithm before being uploaded to the NUWA platform. Therefore, personal information of all included patients in NUWA platform is confidential and unreachable for all users. Several important professionalism levels were also significantly increased in many other elements after the workshop, indicating that this workshop could be a preliminary attempt to promote professionalism when the prominence of digital data has changed

the way we communicate and when doctors have to spend more time with electronic records than with patients. These changes our workshop brought deserve more attention in China since Chinese doctors are well-known to be overburdened.[37, 38] These improvements could also help increase the reliability of medical records and produce convincing and effective medical information.

Another interesting fact in the results is that approximately 60% of students noticed rare diseases in the free-exploration section, which is hardly involved in routine medical classes.[39, 40] We believe this is another strength of big data-based medical education. In traditional medical classes, it is arduous for teachers to grant detailed depictions for every type of rare disease. As a result, students may not have easy access to these exceptional cases and therefore are often not able to diagnose rare diseases when reviewing patients' medical records. The big data platform makes it possible for every user to conveniently browse through cases with rare pathologies, which would benefit patients as well as reduce misdiagnoses.

Our workshop had some limitations. The most important one is that the NUWA platform is still under development, and the omics data are not currently included. Therefore, another investigation should be conducted when the construction of NUWA has been completed. Furthermore, narrow geographical distributions of study participants interfered with the generalization of results to other populations. Meanwhile, there are concerns that students who volunteered to participate in the workshop may be more enthusiastic about big data technology than those who did not. We thus included all 5th- (preclinical) and 6th-grade (clinical) students of 8-year undergraduate education in Tongji Medical School to make the study sample more representative. As for the measurement of attitudes and professionalism, a qualitative approach would have been more appropriate and should be considered in the future. Meanwhile, this workshop is too short to produce fundamental improvements in students' attitudes and professionalism. In the future, more long-term studies are wanted to draw a clear conclusion.

Conclusion

This study depicted Chinese students' knowledge of medical big data for the first time and the NUWA platform-based workshop had potential to improve their understanding of big data and enhance professionalism.

List of abbreviations

P-MEX, Professionalism Mini-Evaluation Exercise instrument. NUWA platform, the National Union of Real-World Gynecologic Oncology Research & Patient Management Platform.

Contributorship

QLG, JHL and XFJ designed the study and survey. QLG, JHL, XFJ, ML, ZKP, YBH and SQZ conducted the workshop. PJ, HYL, JHC, XYL, YY, and SQZ analyzed and interpreted the data. YJZ and GCM performed interpretation of data and discussion of findings. QLG conceptualized and designed the study, supervised the project, analyzed and interpreted the data, and wrote the paper. All authors approved the final version of the manuscript and agreed to the submission of this manuscript.

Funding statement

This work was supported by the Teaching Reform Project for Postgraduates of Huazhong University of Science and Technology (YY202006), the Major Project of Medical Education (20A0769), and Teaching Research Project of The Second Clinical School, Tongji Medical College, Huazhong University of Science and Technology (2021022).

Competing of interests

The authors declare that they have no competing interests.

Ethics approval

The data collection and analysis plan were acknowledged and agreed by all participants and the study has been approved by the Research Ethics Commission of Tongji Hospital of Huazhong University of Science and Technology (2020-S201).

Data sharing

The data used in the current study are available from the corresponding author on reasonable request via e-mail (qingleigao@hotmail.com).

Consent for publication

Not applicable.

Acknowledgements

Not applicable.

Reference

[1] Bezemer T, de Groot MC, Blasse E, et al. A Human(e) Factor in Clinical Decision Support Systems [J]. J Med Internet Res. 2019;21(3):e11732-e.

[2] Sim I. Two Ways of Knowing: Big Data and Evidence-Based Medicine [J]. Ann Intern Med. 2016;164(8):562-3.

[3] Wu WT, Li YJ, Feng AZ, et al. Data mining in clinical big data: the frequently used databases, steps, and methodological models [J]. Military Medical Research. 2021;8(1):44.

[4] Olatosi B, Zhang J, Weissman S, et al. Using big data analytics to improve HIV medical care utilisation in South Carolina: A study protocol [J]. #N/A. 2019;9(7):e027688.

[5] Carra G, Salluh JIF, da Silva Ramos FJ, et al. Data-driven ICU management: Using Big Data and algorithms to improve outcomes [J]. J Crit Care. 2020;60:300-4.

[6] Kumar NMS, Eswari T, Sampath P, et al. Predictive Methodology for Diabetic Data Analysis in Big Data In: Vijayakumar V, Neelanarayanan V, editors. Big Data, Cloud and Computing Challenges [M]. Procedia Computer Science. 502015. p. 203-8.

[7] Razavian N, Blecker S, Schmidt AM, et al. Population-Level Prediction of Type 2 Diabetes From Claims Data and Analysis of Risk Factors [J]. #N/A. 2015;3(4):277-87.

[8] Fitzgerald RC. Big data is crucial to the early detection of cancer [J]. Nat Med. 2020;26(1):19-20.

[9] Fridley BL, Koestler DC, Godwin AK. Individualizing care for ovarian cancer patients using big data [J]. J Natl Cancer Inst. 2014;106(5).

[10] Cammarota G, Ianiro G, Ahern A, et al. Gut microbiome, big data and machine learning to promote precision medicine for cancer [J]. Nat Rev Gastroenterol Hepatol. 2020;17(10):635-

48.
[11] Harmon SA, Sanford TH, Xu S, et al. Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets [J]. Nat Commun. 2020;11(1):4080.
[12] Mei X, Lee HC, Diao KY, et al. Artificial intelligence-enabled rapid diagnosis of patients with COVID-19 [J]. Nat Med. 2020;26(8):1224-8.
[13] Wang CJ, Ng CY, Brook RH. Response to COVID-19 in Taiwan: Big Data Analytics, New Technology, and Proactive Testing [J]. Jama. 2020;323(14):1341-2.
[14] Jensen PB, Jensen LJ, Brunak S. Mining electronic health records: towards better research applications and clinical care [J]. Nature reviews Genetics. 2012;13(6):395-405.
[15] Andreu-Perez J, Poon CC, Merrifield RD, et al. Big data for health [J]. IEEE J Biomed Health Inform. 2015;19(4):1193-208.
[16] O'Connor S, Zhang M, Honey M, et al. Digital professionalism on social media: A narrative review of the medical, nursing, and allied health education literature [J]. Int J Med Inf. 2021;153:104514.
[17] Attai DJ, Anderson PF, Fisch MJ, et al. Risks and benefits of Twitter use by hematologists/oncologists in the era of digital medicine [J]. Seminars in hematology. 2017;54(4):198-204.
[18] Epstein RM, Hundert EM. Defining and assessing professional competence [J]. Jama. 2002;287(2):226-35.
[19] Cain J, Romanelli F. E-professionalism: a new paradigm for a digital age [J]. Currents in Pharmacy Teaching and Learning. 2009;1(2):66-70.
[20] Hao X, Peng X, Ding X, et al. Application of digital education in undergraduate nursing and medical interns during the COVID-19 pandemic: A systematic review [J]. Nurse Educ Today. 2022;108:105183.
[21] Kailey BS, Dev D, Sado DM, et al. CardioWebinar: the evolution of digital education during the COVID-19 pandemic [J]. Heart. 2021;107(24):2004-5.
[22] Cruess R, McIlroy JH, Cruess S, et al. The Professionalism Mini-Evaluation Exercise: A Preliminary Investigation [J]. 2006;81(10):S74-S8.
[23] Cruess R, McIlroy JH, Cruess S, et al. The Professionalism Mini-Evaluation Exercise: A Preliminary Investigation [J]. Academic Medicine. 2006;81(10).
[24] Norcini JJ, Blank LL, Arnold GK, et al. The mini-CEX (clinical evaluation exercise): a preliminary investigation [J]. Ann Intern Med. 1995;123(10):795-9.
[25] Parthiban N, Boland F, Fadil Azim DH, et al. Asian medical students' attitudes towards professionalism [J]. Med Educ Online. 2021;26(1):1927466.
[26] Tsugawa Y, Ohbu S, Cruess R, et al. Introducing the Professionalism Mini-Evaluation Exercise (P-MEX) in Japan: results from a multicenter, cross-sectional study [J]. Acad Med. 2011;86(8):1026-31.
[27] Bajwa NM, Nendaz MR, Posfay-Barbe KM, et al. A Meaningful and Actionable Professionalism Assessment: Validity Evidence for the Professionalism Mini-Evaluation Exercise (P-MEX) Across 8 Years [J]. Acad Med. 2021;96(11s):S151-s7.
[28] Wu J, Wang J, Nicholas S, et al. Application of Big Data Technology for COVID-19 Prevention and Control in China: Lessons and Recommendations [J]. J Med Internet Res. 2020;22(10):e21980.
[29] Shen T, Wang C. Big Data Technology Applications and the Right to Health in China during

the COVID-19 Pandemic [J]. International journal of environmental research and public health. 2021;18(14).

[30] Li X, Cong Y. A Systematic Literature Review of Ethical Challenges Related to Medical and Public Health Data Sharing in China [J]. Journal of Empirical Research on Human Research Ethics. 2021:15562646211040299.

[31] Liang W, Yao J, Chen A, et al. Early triage of critically ill COVID-19 patients using deep learning [J]. Nature communications. 2020;11(1):3543-.

[32] Meo SA, Sattar K, Ullah CH, et al. Progress and prospects of medical education research in Asian Countries [J]. Pakistan journal of medical sciences. 2019;35(6):1475-81.

[33] Schmid RD, Xiong X. Biotech in China 2021, at the beginning of the 14th five-year period ("145") [J]. Applied microbiology and biotechnology. 2021;105(10):3971-85.

[34] Wu H, Li W, Xie A, et al. Funding of health professional education: China's 20-year process and a global comparison [J]. Medical education. 2021.

[35] Silva Junior G, Askari M, Oliveira J. Digital Health and Possible Solutions to Improve the Care in the Field of Nephrology [J]. Contributions to nephrology. 2021;199:1-15.

[36] Doukas DJ, McCullough LB, Wear S. Perspective: Medical education in medical ethics and humanities as the foundation for developing medical professionalism [J]. Acad Med. 2012;87(3):334-41.

[37] Liu S, Chen Y, Li S, et al. What Are the Important Factors Influencing the Recruitment and Retention of Doctoral Students in a Public Health Setting? A Discrete Choice Experiment Survey in China [J]. International journal of environmental research and public health. 2021;18(18).

[38] Jingang A. Which future for doctors in China? [J]. Lancet. 2013;382(9896):936-7.

[39] Walkowiak D, Domaradzki J. Are rare diseases overlooked by medical education? Awareness of rare diseases among physicians in Poland: an explanatory study [J]. Orphanet journal of rare diseases. 2021;16(1):400.

[40] Domaradzki J, Walkowiak D. Knowledge and Attitudes of Future Healthcare Professionals Toward Rare Diseases [J]. Front Genet. 2021;12:639610.

Figure Legends

Figure 1. Basic knowledge for big data platform (A) before and (B) after the workshop.

* means that there is a significant difference before and after the workshop.

Figure 2. Students' attitude towards big data platform (A) before and (B) after the workshop.

* means that there is a significant difference before and after the workshop.

Figure 3. The professionalism for students (A) before and (B) after the workshop.

* means that there is a significant difference before and after the workshop.

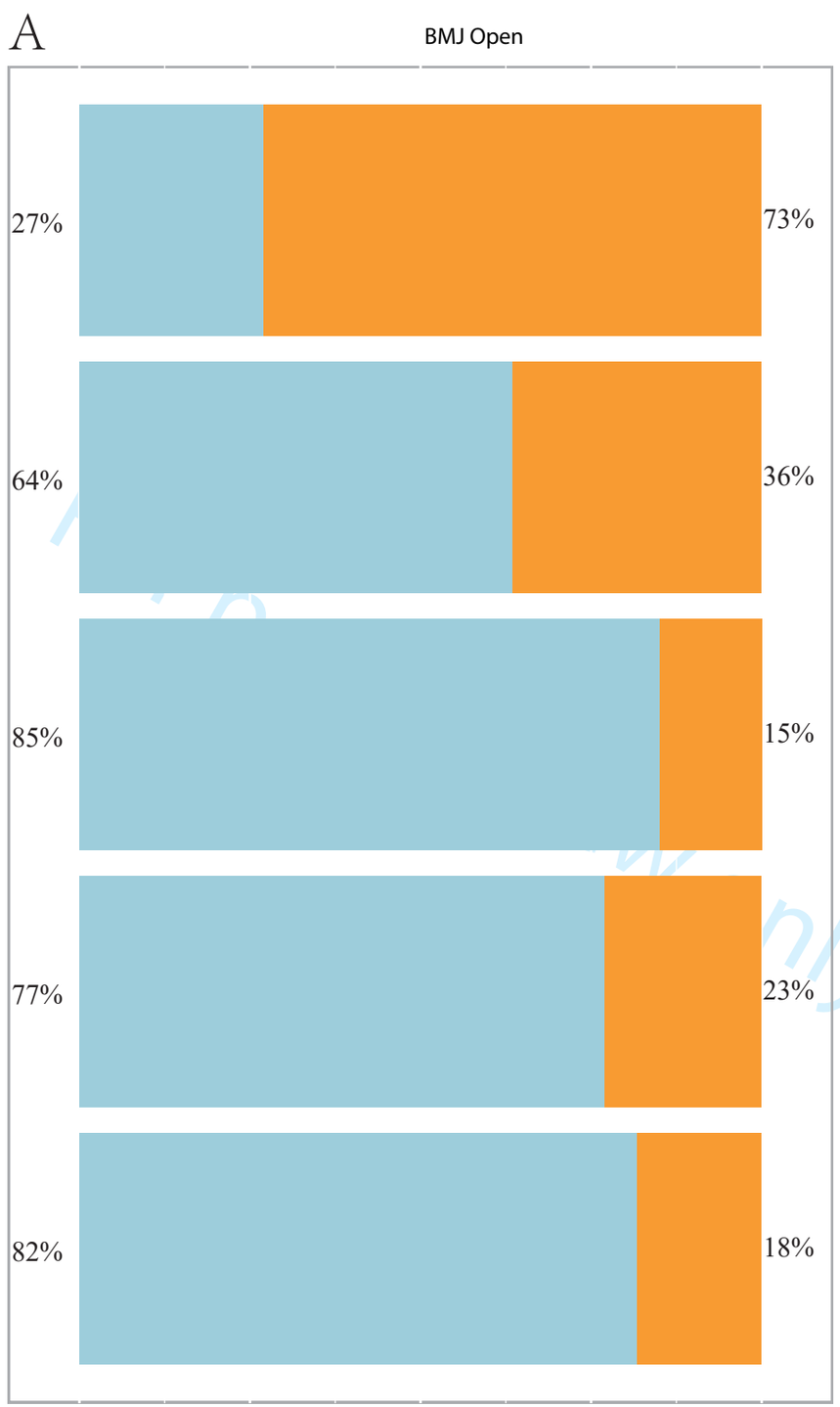
*I know the content of big data platform

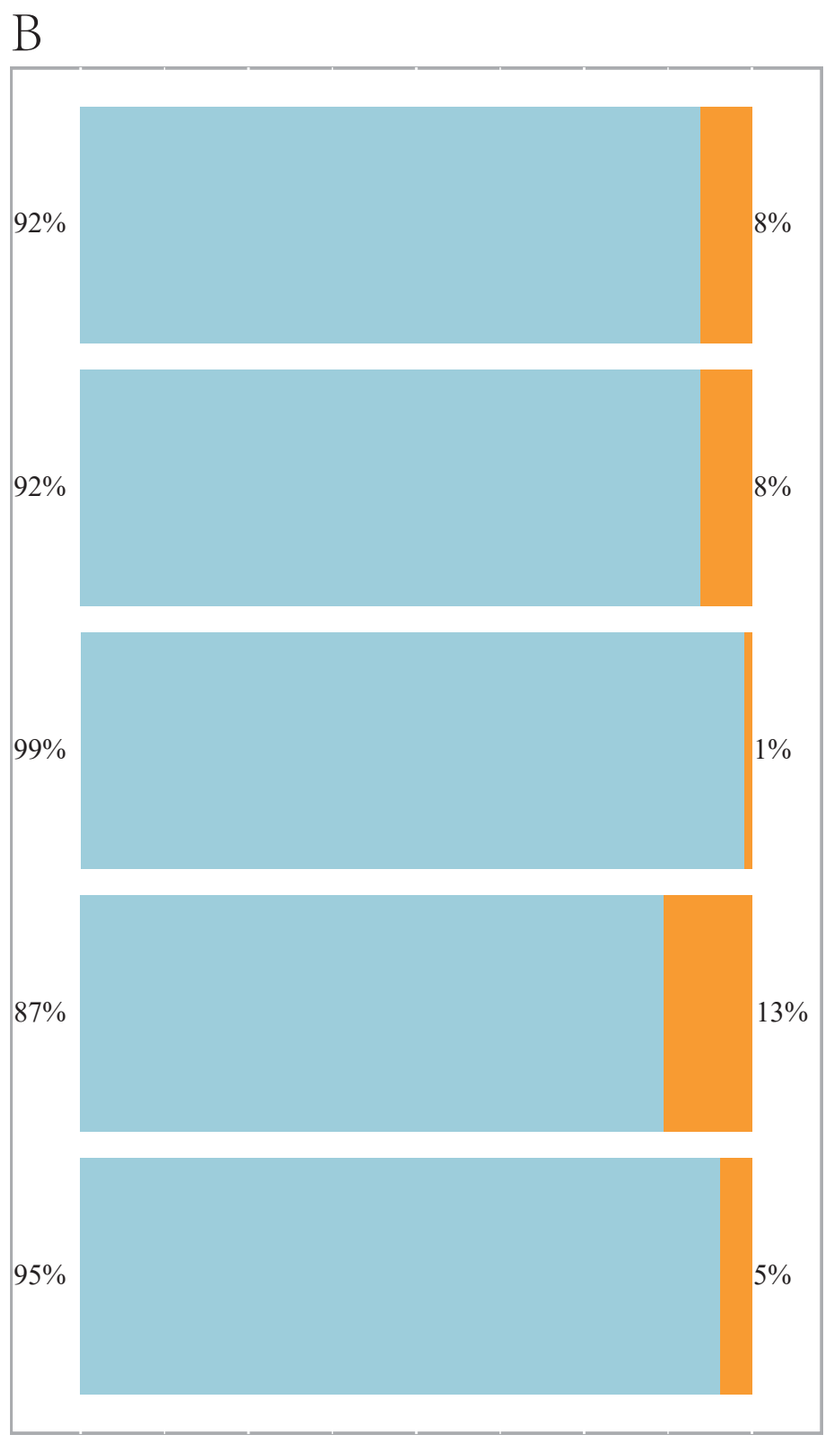
*I know the potential applications of big data technology

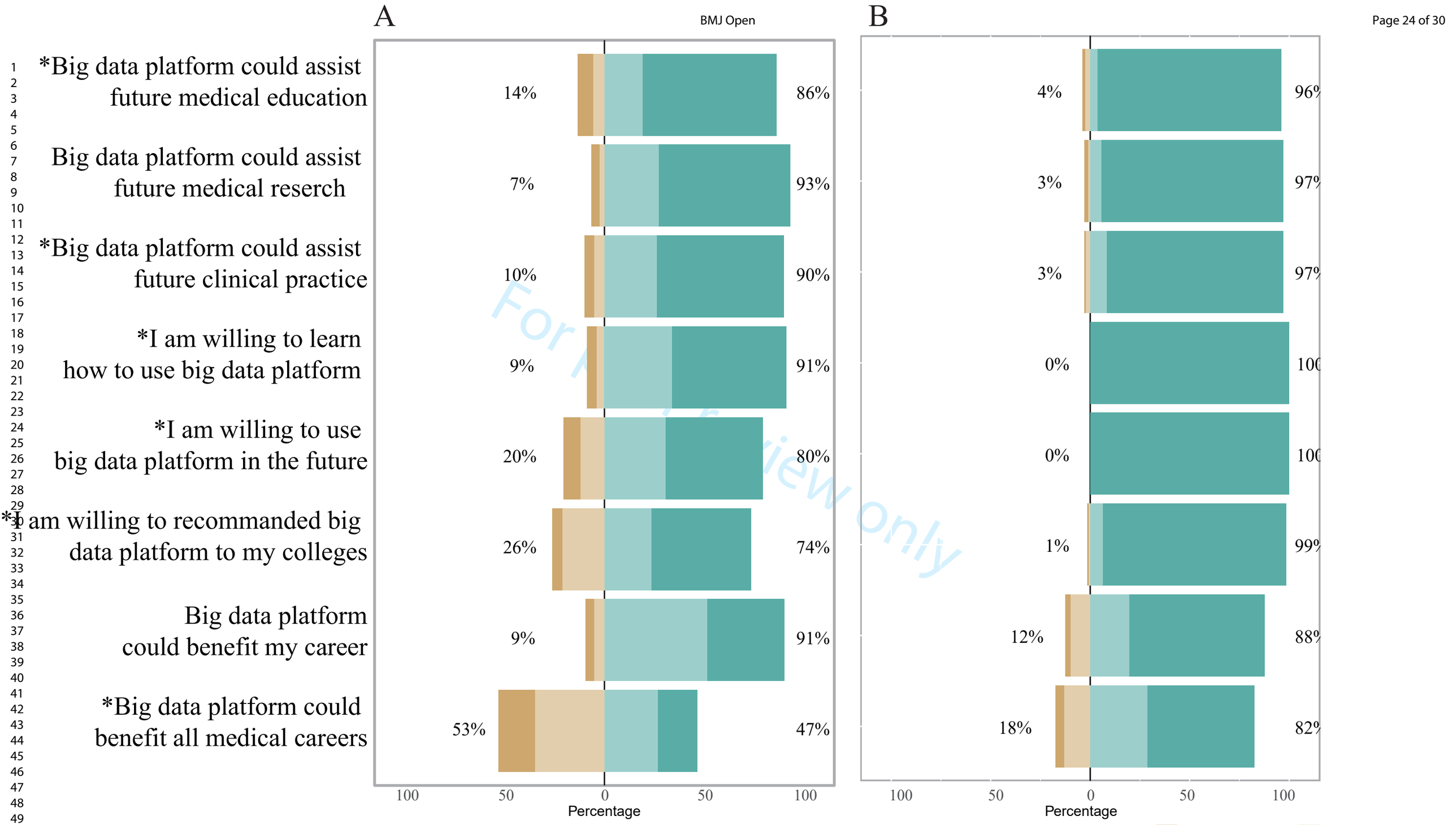
*Big data technology is practical in medical education

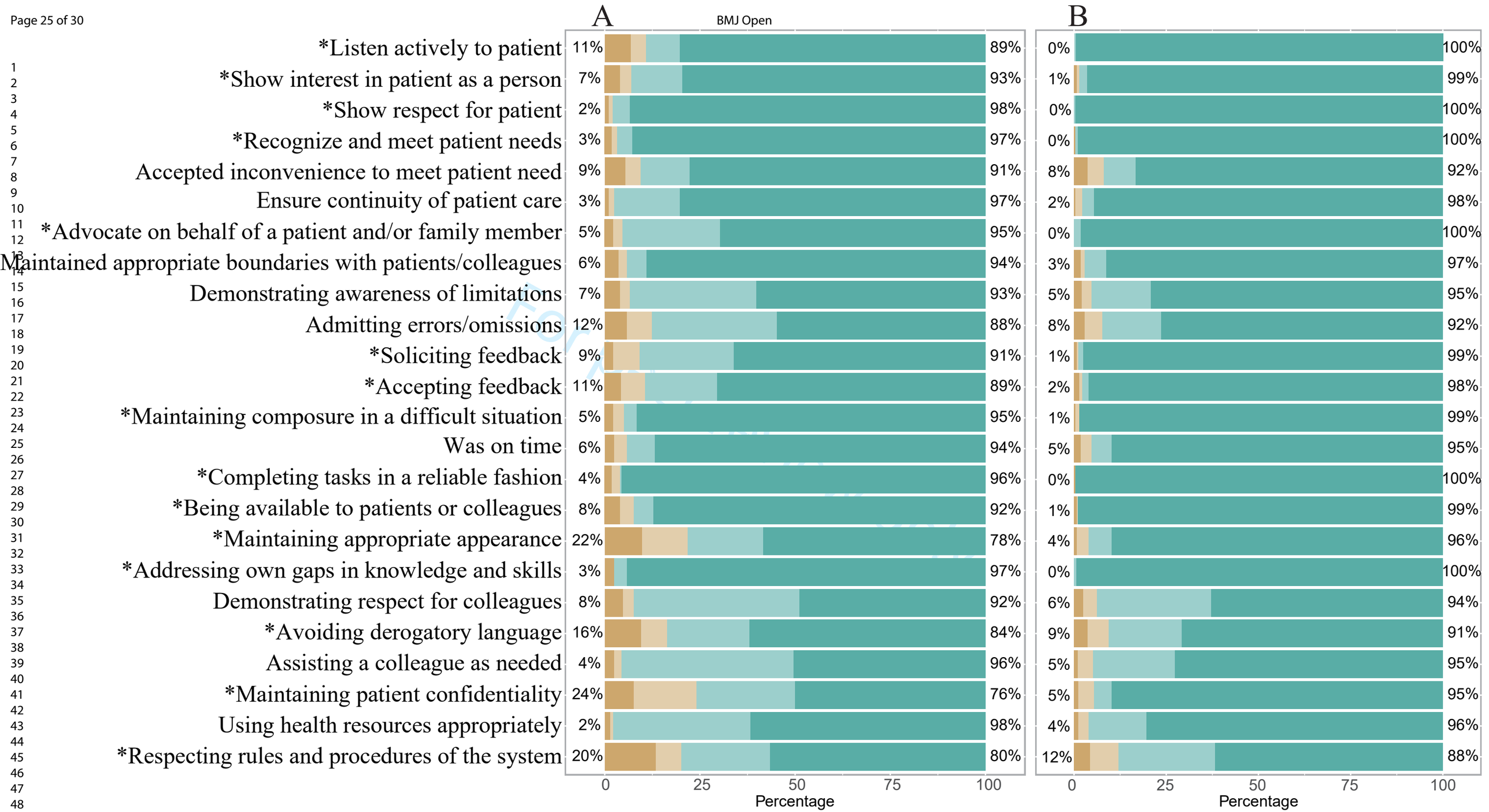
*Big data technology is practical in clinical practice

*Big data technology is practical in scientific research









Supplemental tables

Table S1, basic knowledge for big data platform before and after the workshop.

	Before Workshop		After Workshop		
Items	YES, n	NO, n	YES, n	NO, n	*P*
I know the content of big data platform	74	200	253	21	*<0.0001*
I know the potential applications of big data technology	174	100	253	21	*<0.0001*
Big data technology is practical in medical education	233	41	271	3	*<0.0001*
Big data technology is practical in clinical practice	211	63	238	36	*0.0039*
Big data technology is practical in scientific research	224	50	261	13	*<0.0001*

Table S2, students' attitude towards big data platform before and after the workshop.

	Before Workshop						After Workshop						*P*
Items	Totally disagree, n	Disagree, n	Agree, n	Totally agree, n	Mean	SD	Totally disagree, n	Disagree, n	Agree, n	Totally agree, n	Mean	SD	
Big data platform could assist future medical education	21	16	53	184	3.50	0.910	4	7	10	253	3.87	0.504	*0.0002*
Big data platform could assist future scientific research	11	7	75	181	3.55	0.735	5	3	16	250	3.86	0.499	*0.0705*
Big data platform could assist future clinical practice	13	14	72	175	3.49	0.799	2	6	23	243	3.85	0.464	*0.0017*
I am willing to learn how to use big data platform	13	11	93	157	3.43	0.783	0	0	0	274	4	0	*<0.0001*

I am willing to use big data platform in the future	23	33	84	134	3.20	0.953	0	0	0	274	4	0	*<0.0001*
I am willing to recommended big data platform to my colleagues	14	58	65	137	3.19	0.941	2	2	18	252	3.90	0.388	*<0.0001*
Big data platform could benefit my career	12	14	142	106	3.25	0.744	7	27	54	186	3.53	0.776	*0.3382*
Big data platform could benefit all medical careers	50	96	74	54	2.48	1.006	12	36	79	147	3.32	0.863	*<0.0001*

SD, Standard Deviation.

Table S3, the professionalism for students before and after the workshop.

	Before workshop				After workshop				*P*
Items	Not at all important, n	Not important, n	Important, n	Very important, n	Not at all important, n	Not important, n	Important, n	Very important, n	
Doctor-patient relationship skills									
Listen actively to patient	19	11	24	220	0	0	1	273	*<0.0001*
Show interest in patient as a person	11	8	37	218	2	2	6	264	*0.0029*
Show respect for patient	3	3	12	256	0	0	1	273	*0.0401*
Recognize and meet patient needs	5	4	11	254	1	0	2	271	*0.0255*
Accepted inconvenience to meet patient need	15	11	35	213	10	12	24	228	*0.6503*
Ensure continuity of patient care	3	4	47	220	1	5	9	259	*1*

Advocate on behalf of a patient and/or family member	6	7	70	191	0	0	5	269	*0.0008*
Maintained appropriate boundaries with patients/colleagues	10	6	14	244	5	3	16	250	*0.1440*
Reflective skills									
Demonstrating awareness of limitations	11	7	91	165	6	7	44	217	*0.4595*
Admitting errors/omissions	16	18	90	150	8	13	44	209	*0.0880*
Soliciting feedback	6	19	68	181	2	1	4	267	*<0.0001*
Accepting feedback	12	17	52	193	4	2	5	263	*0.0001*
Maintaining composure in a difficult situation	6	8	9	251	1	3	0	270	*0.0310*
Time management									
Was on time	7	9	20	238	5	8	15	246	*0.7027*
Completing tasks in a reliable fashion	5	6	1	262	1	0	0	273	*0.0086*
Being available to patients or colleagues	11	10	14	239	2	1	0	271	*0.0004*
Interprofessional relationship skills									
Maintaining appropriate appearance	27	33	54	160	2	9	17	246	*<0.0001*
Addressing own gaps in knowledge and skills	7	0	9	258	0	0	2	272	*0.0225*
Demonstrating respect for colleagues	13	8	119	134	7	10	85	172	*0.6139*
Avoiding derogatory language	26	19	59	170	10	16	54	194	*0.0220*
Assisting a colleague as needed	7	5	124	138	3	11	61	199	*0.8407*

Maintaining patient confidentiality	21	45	71	137	3	12	13	246	*<0.0001*
Using health resources appropriately	4	2	99	169	3	8	43	220	*0.3244*
Respecting rules and procedures of the system	37	18	64	155	12	21	72	169	*0.0146*

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4
Objectives	3	State specific objectives, including any prespecified hypotheses	4-5
Methods			
Study design	4	Present key elements of study design early in the paper	5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	5
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case	NA
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	6-7
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	6-7
Bias	9	Describe any efforts to address potential sources of bias	NA
Study size	10	Explain how the study size was arrived at	6
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	7-8
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	7-8
		(*b*) Describe any methods used to examine subgroups and interactions	NA
		(*c*) Explain how missing data were addressed	7-8
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	7-8
		(*e*) Describe any sensitivity analyses	NA

Continued on next page

Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	8
		(b) Give reasons for non-participation at each stage	All participants were eligible and included in this study.
		(c) Consider use of a flow diagram	NA
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	8-9
		(b) Indicate number of participants with missing data for each variable of interest	Very high response rates with no missing data, as shown in supplementary data.
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	NA
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	NA
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	NA
		Cross-sectional study—Report numbers of outcome events or summary measures	8
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	In all tables and figures.
		(*b*) Report category boundaries when continuous variables were categorized	In all tables and figures.
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	NA
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	NA
Discussion			
Key results	18	Summarise key results with reference to study objectives	13
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	15
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	15-16
Generalisability	21	Discuss the generalisability (external validity) of the study results	15-16
Other information			

Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	18

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

Oncologic big data platform for promoting digital competencies and professionalism in Chinese medical students

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2022-061015.R2
Article Type:	Original research
Date Submitted by the Author:	18-Aug-2022
Complete List of Authors:	Liu, Jiahao; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Jiao, Xiaofei; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Zeng, Shaoqing; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Li, Huayi; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Jin, Ping; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Chi, Jianhua; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Liu, Xingyu; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Yu, Yang; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Ma, Guanchen; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Zhao, Yingjun; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Li, Ming; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Peng, Zikun; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Huo, Yabing; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Gao, Qing-Lei ; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology
Primary Subject Heading:	Medical education and training
Secondary Subject Heading:	Medical education and training
Keywords:	MEDICAL EDUCATION & TRAINING, Gynaecological oncology < GYNAECOLOGY, Gynaecological oncology < ONCOLOGY, Health informatics < BIOTECHNOLOGY & BIOINFORMATICS

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Oncologic big data platform for promoting digital competencies and professionalism in Chinese medical students

Jiahao Liu[1,2#], Xiaofei Jiao[1,2#], Shaoqing Zeng[1,2], Huayi Li[1,2], Ping Jin[1,2], Jianhua Chi[1,2], Xingyu Liu[1,2], Yang Yu[1,2], Guanchen Ma[1,2], Yingjun Zhao[1,2], Ming Li[1,2], Zikun Peng[1,2], Yabing Huo[1,2], Qinglei Gao[1,2*]

[1] Department of Obstetrics and Gynecology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, Hubei, P. R. China.
[2] Cancer Biology Research Center (Key Laboratory of the Ministry of Education), Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, Hubei, P. R. China.

[#] These authors contribute equally.

* Correspondence to:
Prof. Qinglei Gao, Cancer Biology Research Center (Key Laboratory of the Ministry of Education), Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, 1095 Jiefang Ave, Wuhan 430030, Hubei, P. R. China. Tel: +86-27-83662681; Fax: +86-27-83662681; E-mail: qingleigao@hotmail.com.

Word count (excluding title page, abstract, references, figures and tables): 3,198

Abstract

Objectives: Advancements in big data technology are reshaping the health care system in China. This study aims to explore the role of medical big data in promoting digital competencies and professionalism among Chinese medical students.

Design, setting and participants: This study was conducted among 274 medical students who attended a workshop on medical big data conducted on July 8, 2021 in Tongji Hospital. The workshop was based on the first nationwide multifunction gynecologic oncology medical big data platform in China, at the National Union of Real-World Gynecologic Oncology Research & Patient Management Platform (NUWA platform).

Outcome measures: Data on knowledge, attitudes toward big data technology and professionalism were collected before and after the workshop. We have measured the four skill categories: doctor–patient relationship skills, reflective skills, time management, and interprofessional relationship skills using the Professionalism Mini–Evaluation Exercise (P-MEX) as a reflection for professionalism.

Results: A total of 274 students participated in this workshop and completed all the surveys. Before the workshop, only 27% of them knew the detailed content of medical big data platforms, and 64% knew the potential application of medical big data. The majority of the students believed that big data technology is practical in their clinical practice (77%), medical education (85%), and scientific research (82%). Over 80% of the participants showed positive attitudes toward big data platforms. They also exhibited sufficient professionalism before the workshop. Meanwhile, the workshop significantly promoted students' knowledge of medical big data ($P < 0.05$), and led to more positive attitudes toward big data platforms and higher levels of professionalism.

Conclusions: Chinese medical students have primitive acquaintance and positive attitudes toward big data technologies. The NUWA platform-based workshop may potentially promote their understanding of big data and enhance professionalism, according to the self-measured P-MEX scale.

Keywords: Big data, Chinese medical education, Digital competencies,

Professionalism, Workshop

Strengths and limitations of this study

- This study depicted the knowledge and professionalism of medical big data among Chinese medical students in the era of big data for the first time.
- This study provided a vivid example of how big data can assist medical education, based on the first nationwide gynecologic oncology medical big data platform in China (NUWA platform).
- The generalization of the findings is limited by the sample representation. The method of measuring attitudes and professionalism may be improved in the future. We need longer time to observe the change of digital competencies and professionalism in Chinese medical students.

Background

Electronification of medical records is the signature of the modern health care system.[1-3] Massive clinical and omics data have been produced to enable more detailed depictions of patients and diseases. Medical big data is thus reshaping our appreciation of the modern medical system. Technological advancements in data storage, processing, and analysis accelerated the clinical application of big data-driven products and contributed to personalized disease management,[4, 5] early diagnosis,[6-8] and treatment decision.[9, 10] Especially in the COVID-19 pandemic setting, achieving rapid application of medical big data would meet the pressing clinical need to predict the progression of diseases using data characteristics.[11-13]

Big data has brought new challenges for doctors.[14] Information overload is a challenge for every healthcare worker, since they have to acclimate to the nature of big data, including extraordinary value, volume, velocity, variety, and variability.[15] Challenges can also be expected regarding medical professionalism in the age of big data. Because big data values realistic health-related information more than ever, the illusion that digital data outweigh face-to-face physician–patient interactions may motivate doctors to ignore the importance of professionalism.[16, 17] There is no precise definition of medical professionalism, which is reflected in the attitudes and behaviors directly related to clinical practice. Epstein RM et al. proposed the definition of professional competence as wisely using communication, knowledge, technical skills, clinical reasoning, emotions, values, and reflection in daily practice.[18] And Cain J et al. summarized professionalism as a series of attributes, such as altruism, respect, honesty, and so on.[19] Some studies have shown that digital medical education based on big data plays a positive effect in promoting professionalism.[16, 20, 21] However, there is a lack of research on the attitude and professionalism of Chinese healthcare workers regarding medical big data. Therefore, extra lectures and workshops about obtaining insights into big data and remaining respectful to patients are necessary.

In this study, we aimed at depicting the knowledge of medical big data and professionalism in the era of big data among Chinese medical students. What's more,

we conducted a workshop for Chinese clinical and preclinical students to improve their knowledge of medical big data and their professionalism in the big data era, based on the National Union of Real-World Gynecologic Oncology Research & Patient Management Platform (NUWA platform), which is the first nationwide multifunction gynecologic oncology medical big data platform in China.

Methods

Study design

A workshop on the introduction of medical big data was conducted among clinical and preclinical students in Tongji Medical School. We conducted this study on Chinese medical students' learning and the application of big data in healthcare using pre- and post-course questionnaires.

Workshop design

The workshop was conducted on July 8, 2021 in Tongji Hospital, Tongji Medical School, Huazhong University of Science and Technology, Wuhan, China. It was part of the series class of "Medical big data platform learning and applying".

The workshop was delivered by two senior doctors (QLG and SQZ), who have over 5 years of experience in medical big data platform development and application. The workshop consisted of three major parts, (i) a lecture on the content and application of medical big data; (ii) a lecture on professionalism for doctors in the big data era; and (iii) learners being allowed to explore the NUWA platform freely according to their interest. The first and second sections lasted for approximately 45 minutes, and the third section lasted for 30 minutes.

In section one, five major parts were included in lectures given by the senior doctors, (i) the development of medical big data in China and the world, (ii) the application and potential of big data in clinical practice, medical education, and scientific research, (iii) the construction and content of the NUWA platform, (iv) how to use the NUWA

platform, and (v) plans for the NUWA platform development. For the second section, the following four items were discussed: (i) the attitude health care workers should have when communicating with patients and colleagues considering that big data have already reshaped our medical system, (ii) the importance of detailed and continuous patient information for the development of big data platforms, (iii) how to protect patient privacy on the databases, and (iv) how to use the NUWA platform to satisfy patients' needs. For Section 3, all students were given a temporary account for the NUWA platform and have access to all deidentified patient information. They were allowed to explore the medical data freely and view structured health care information for half an hour.

Data collection

Clinical and preclinical students in Tongji Medical School were invited to attend a 2-hour class by email or roadshows between June 8, 2021, and July 7, 2021. The contents and speakers for the workshop were presented in the email or during the roadshow. In addition, all 5th- (n=50, preclinical) and 6th-grade (n=49, clinical) students of 8-year undergraduate education in Tongji Medical School took part in this workshop as an additional course.

The participants completed two surveys, one before the workshop and the other after the workshop. Both questionnaires contain their basic knowledge and attitude toward big data technology and a Professionalism Mini-Evaluation Exercise (P-MEX) instrument. The P-MEX is developed from the mini-Clinical Examination Exercise (mini-CEX) format by Cruess R et al. in 2006 to evaluate professionalism in clinical training.[22] It consists of 24 items representing four skill categories: doctor–patient relationship skills, reflective skills, time management, and interprofessional relationship skills skills.[23, 24] And the reliability and validity of P-MEX have been confirmed in both Eastern and Western cultural backgrounds.[25-27] The attitude scale is a self-created scale focused on measuring participants' pedagogic evaluation, acceptance and expectation of the big data platform. It consisted of 8 questions: (i) Big

data platform could assist future medical education, (ii) Big data platform could assist future medical research, (iii)Big data platform could assist future clinical practice, (iv) I am willing to learn how to use big data platform, (v) I am willing to use big data platform in the future, (vi) I am willing to recommend big data platform to my colleagues, (vii) Big data platform could benefit my career, and (viii) Big data platform could benefit all medical careers. For each question, students chose from "totally agree", "agree", "disagree", and "totally disagree". For the applicability of conclusions, "totally agree" and "agree" were regarded as "positive attitude", and "disagree" and "totally disagree" were considered as "negative attitude". In addition, the first survey also contained students' baseline information, and the second survey contained a satisfaction questionnaire. Meanwhile, the browsing histories for participants in Section 3 were also recorded and analyzed to reflect students' interest.

The data collection and analysis plan were acknowledged and agreed upon by all participants at the beginning of the workshop, and the study was approved by the Research Ethics Commission of Tongji Hospital of Huazhong University of Science and Technology (2020-S201).

NUWA platform

The NUWA platform is the first nationwide Gynecological Oncology data-sharing platform launched by the National Clinical Research Center for Gynecological Oncology in August 2019. This platform integrated inpatient/outpatient clinical data, genomic data, and follow-up data to develop a patient-level longitudinal clinico-genomic database. Information was deidentified and extracted from electronic medical records. A rigorous data quality check was performed to ensure the accuracy of the data entries. Since its foundation in 2019, 17 first-class hospitals from different provinces or cities in China have participated in the NUWA platform until August 2021.

Statistical analysis

Descriptive statistics were presented by counts and percentages to describe the

demographic information. The Chi-square test was used to compare the changes in knowledge and understanding of big data and professionalism before and after the workshop. To acquire more practical results, the categories "not at all important" and "not important" were combined for the analysis, as well as "important" and "very important". *P values <0.05* were considered statistically significant. The data were analyzed using R Version 4.03.

Patient and Public Involvement

No patients or members of the public were involved in this study.

Results

Characteristics of students

A total of 274 students participated in this workshop and completed two surveys. All of them were included in the final analysis. Participants were aged between 22 and 28, with 148 (54.0%) preclinical students and 126 (46.0%) clinical students. Among them, 130 (47.4%) were males and 144 (52.4%) were females (**Table 1**).

The majority of participants (207, 75.5%) knew of at least one big data platform in China or the world. Over two-thirds of them (183, 66.8%) also acknowledged its application. However, only 17.2% (47) of them had been involved in any project related to medical big data (**Table 1**).

Table 1. Baseline characteristics for participants.

Characteristics	Number of participants, n (%)
Age, years	
22	27 (9.9%)
23	16 (5.8%)
24	67 (24.5%)
25	66 (24.1%)
26	50 (18.2%)

	27	28 (10.2%)
	28	20 (7.3%)
Gender		
	male	130 (47.4%)
	female	144 (52.6%)
Study stage		
	pre-clinical	148 (54.0%)
	clinical	126 (46.0%)
Acknowledgment of any kind of big data platform		
	yes	207 (75.5%)
	no	67 (24.5%)
Know the applications of big data technology		
	yes	183 (66.8%)
	no	91 (33.2%)
Involved in any big data-related projects		
	yes	47 (17.2%)
	no	227 (82.8%)
Total		274 (100%)

Knowledge of big data platforms

Before the workshop, approximately a quarter of the students (74, 27%) knew detailed content of at least one medical big data platform, while nearly 64% (174) knew the potential application of medical big data platforms. After attending the lessons, almost all students could understand the content (253, 92%) and the potential application (253, 92%) of medical big data (**Figure 1 and Table S1**).

Regardless of the survey completed (pre- or post-course), the majority of students

believed that big data technology is practical in medical education, clinical practice, and scientific research (85%, 77%, 82% before the workshop; 99%, 87%, 95% after the workshop, respectively) (**Figure 1 and Table S1**).

Students' attitudes toward big data platforms

Even before the workshop, most students hold a positive attitude on the potential of the big data platform in promoting medical education (237, 86%), medical research (256, 93%), and clinical practices (247, 90%). They were also willing to learn about and use the big data platform (250, 91%, and 218, 80% respectively). Of them, 74% (202) were enthusiastic to introduce big data platforms to their colleagues. Interestingly, most students were convinced that the big data platform could benefit their careers (248, 91%), but they were not sure if it would yield the same effect on the others (128, 47%) (**Figure 2 and Table S2**).

When the workshop was completed, almost all students had positive attitudes toward the big data platform (**Figure 2 and Table S2**). However, there were still 12% (34) and 18% (48) of students who were not sure if the big data platform could benefit their and others' medical careers, respectively.

Professionalism

Generally, positive attitudes toward all the professionalism items were demonstrated, with more than three-quarters of the students agreeing that all professionalism attributes were "important" or "very important" before the workshop. The three items that most students thought were "not at all important" or "not important" were maintaining patient confidentiality (66, 24% of students chose "not important" or "not at all important"), maintaining appropriate appearance (60, 22% students chose "not important" or "not at all important") and respecting rules and procedures of the system (55, 20% students chose "not important" or "not at all important") (**Figure 3A and Table S3**).

After the lectures, students exhibited a more positive attitude on most

professionalism items ($p<0.05$). The most significant improvements occurred for "Maintaining patient confidentiality" (from 76% to 95%), "Listen actively to patient" (from 89% to 100%), and "Accepting feedback" (from 89% to 98%) (**Figure 3B and Table S3**).

Interest in medical data

In the free exploration section, most students (253, 92.3%) viewed the patients' hospitalization logs. A total of 76.3% (209) of students were interested in the patients' history of illness. Meanwhile, many students were interested in medicine usages (56.2%), surgery reports (52.9%), and diseases of rare pathologies (58.6%). Only approximately one-third of students viewed image diagnosis (37.6%) and follow-up records (32.5%) (**Table 2**).

Table 2. Browser records in the free-exploration section.

Content	Number of participants, n (%)
History of Illness	
yes	209 (76.3%)
no	65 (23.7%)
Hospitalization logs	
yes	253 (92.3%)
no	21 (7.7%)
Medicine usage	
yes	154 (56.2%)
no	120 (43.8%)
Surgery record	
yes	145 (52.9%)
no	129 (47.1%)
Image diagnosis	
yes	103 (37.6%)

no	171 (62.4%)
Follow-up records	
yes	89 (32.5%)
no	185 (67.5%)
Rare pathologies	
yes	160 (58.4%)
no	114 (41.6%)
Total	274 (100%)

Satisfactory survey

Upon completing the survey, 95.6% (262) of participants were "satisfied" or "extremely satisfied" with this workshop, and only 7 out of 274 students were not at all satisfied. However, we failed to obtain feedback from them in the next three months. Most students thought that the workshop was informative (249, 90.9%) and understandable (255, 94.1%). The majority of students were also willing to recommend this seminar to other students (257, 93.8%) and to participate in similar classes in the future (254, 92.7%). Regarding the duration of this workshop, only 7 (2.6%) students thought that 2 hours was too long (**Table 3**).

Table 3. Participants 'answers to the workshop satisfaction survey

	Not at all, n (%)	No, n (%)	Yes, n (%)	Yes, extremely, n (%)	Mean	SD
Overall, are you satisfied with this course	7 (2.6%)	5 (1.8%)	57 (20.8%)	205 (74.8%)	3.68	0.640
Did you think the course are informative?	10 (3.6%)	15 (5.5%)	101 (36.9%)	148 (54%)	3.41	0.757
Did you think the duration of this courses is too long	121 (44.2%)	118 (43.1%)	28 (10.2%)	7 (2.6%)	2.07	0.954

Was the course understandable for you	11 (4%)	8 (2.9%)	121 (44.2%)	134 (48.9%)	3.38	0.733
Would you recommend these courses to other students?	6 (2.2%)	11 (4%)	144 (52.6%)	113 (41.2%)	3.33	0.659
Are you willing to take part in similar courses in the future?	6 (2.2%)	14 (5.1%)	35 (12.8%)	219 (79.9%)	3.70	0.666

SD, Standard Deviation.

Discussion

Our study demonstrated that Chinese medical students have little knowledge of and positive attitudes toward big data technologies. They also yield expertise in professionalism. Furthermore, workshops based on big data platforms could further strengthen their digital competencies and improve doctor–patient communication capabilities, which would lead to better fitness during the expansion of medical big data.

In this workshop, the basic knowledge of big data platforms in China was measured. Although a large fraction of students knew about the big data platform and its utility, only a few of them could apprehend its composition or how to use it. This situation may be caused by the recent rapid development of big data technology in China and comparatively lagging relevant education.[28, 29] In addition, ethical challenges that hinder medical and public health data sharing may have also exacerbated the situation.[30] Therefore, aside from conducting big data-related workshops, promoting medical and public health data sharing policies and evolving relevant legal and ethical implications are also of great importance.

Students' enthusiasm toward medical big data was also taken into consideration. As expected and consistent with previous reports,[25] the majority of students exhibited positive attitudes toward the big data platform and were willing to acquire further

proficiency. They believed that big data would play a vital role in future medical education, clinical practice, and scientific research, which are the main tasks for all Chinese doctors. However, not all of them believed that big data could benefit all medical careers. The popularity of information technology of China recently may contribute to the big data-friendly intention in medical students.[31] Meanwhile, their limited understanding of related fields may hamper the belief that it could be used productively.[32-34] In the meantime, we were delighted to see that almost all students recognized the significance of big data after our workshop. We thus believe that proper education may be an efficient way to diminish misunderstandings and to achieve full potential in big data platforms.

Professionalism was related to the development of big data in our study for the first time in China. Considering that big data can reshape medical activities in all aspects, including doctors' attitudes toward their colleagues and patients,[35] a lecture about "professionalism in the age of big data" was presented in the current workshop. Before the seminar, students demonstrated sufficient professionalism similar to the findings of previous reports, with high scores on the majority of these items. Nevertheless, it is worth noting that nearly 1/4 of the participants did not pay enough attention to maintaining patient confidentiality. Protecting patients' privacy is an essential embodiment of medical ethics and humanities.[36] Participants mostly answered these questions from the perspective of big data users before our workshop, ignoring that those data represented thousands of actual patients. After we emphasized the importance of data privacy in the workshop, students realized that respect for patients is the foundation for improving their medical professionalism. Meanwhile, to enhance patient-data privacy protection, the privacy information, such as ID number, was privatized and converted into an alternative ID number using a hashing algorithm before being uploaded to the NUWA platform. Therefore, personal information of all included patients in NUWA platform is confidential and unreachable for all users. Several important professionalism levels were also significantly increased in many other elements after the workshop, indicating that this workshop could be a preliminary

attempt to promote professionalism when the prominence of digital data has changed the way we communicate and when doctors have to spend more time with electronic records than with patients. These changes our workshop brought deserve more attention in China since Chinese doctors are well-known to be overburdened.[37, 38] These improvements could also help increase the reliability of medical records and produce convincing and effective medical information.

Another interesting fact in the results is that approximately 60% of students noticed rare diseases in the free-exploration section, which is hardly involved in routine medical classes.[39, 40] We believe this is another strength of big data-based medical education. In traditional medical classes, it is arduous for teachers to grant detailed depictions for every type of rare disease. As a result, students may not have easy access to these exceptional cases and therefore are often not able to diagnose rare diseases when reviewing patients' medical records. The big data platform makes it possible for every user to conveniently browse through cases with rare pathologies, which would benefit patients as well as reduce misdiagnoses.

Our workshop had some limitations. The most important one is that the NUWA platform is still under development, and the omics data are not currently included. Therefore, another investigation should be conducted when the construction of NUWA has been completed. Meanwhile, we did not calculate the sample size before the conduction of this study, which may lead to potential bias. And the limited sample size may be another reason that restricts the popularization of our conclusion. Furthermore, narrow geographical distributions of study participants interfered with the generalization of results to other populations. Meanwhile, there are concerns that students who volunteered to participate in the workshop may be more enthusiastic about big data technology than those who did not. We thus included all 5th- (preclinical) and 6th-grade (clinical) students of 8-year undergraduate education in Tongji Medical School to make the study sample more representative. As for the measurement of attitudes and professionalism, a qualitative approach would have been more appropriate and should be considered in the future. Meanwhile, this workshop is too short to

produce fundamental improvements in students' attitudes and professionalism. In the future, more long-term studies are wanted to draw a clear conclusion.

Conclusion

This study depicted Chinese students' knowledge of medical big data for the first time and the NUWA platform-based workshop had potential to improve their understanding of big data and enhance professionalism.

List of abbreviations

P-MEX, Professionalism Mini-Evaluation Exercise instrument. NUWA platform, the National Union of Real-World Gynecologic Oncology Research & Patient Management Platform.

Contributorship

QLG, JHL and XFJ designed the study and survey. QLG, JHL, XFJ, ML, ZKP, YBH and SQZ conducted the workshop. PJ, HYL, JHC, XYL, YY, and SQZ analyzed and interpreted the data. YJZ and GCM performed interpretation of data and discussion of findings. QLG conceptualized and designed the study, supervised the project, analyzed and interpreted the data, and wrote the paper. All authors approved the final version of the manuscript and agreed to the submission of this manuscript.

Funding statement

This work was supported by the Teaching Reform Project for Postgraduates of Huazhong University of Science and Technology (YY202006), the Major Project of Medical Education (20A0769), and Teaching Research Project of The Second Clinical School, Tongji Medical College, Huazhong University of Science and Technology (2021022).

Competing of interests

The authors declare that they have no competing interests.

Ethics approval

The data collection and analysis plan were acknowledged and agreed by all participants and the study has been approved by the Research Ethics Commission of Tongji Hospital of Huazhong University of Science and Technology (2020-S201).

Data sharing

The data used in the current study are available from the corresponding author on reasonable request via e-mail (qingleigao@hotmail.com).

Consent for publication

Not applicable.

Acknowledgements

Not applicable.

Reference

[1] Bezemer T, de Groot MC, Blasse E, et al. A Human(e) Factor in Clinical Decision Support Systems [J]. J Med Internet Res. 2019;21(3):e11732-e.

[2] Sim I. Two Ways of Knowing: Big Data and Evidence-Based Medicine [J]. Ann Intern Med. 2016;164(8):562-3.

[3] Wu WT, Li YJ, Feng AZ, et al. Data mining in clinical big data: the frequently used databases, steps, and methodological models [J]. Military Medical Research. 2021;8(1):44.

[4] Olatosi B, Zhang J, Weissman S, et al. Using big data analytics to improve HIV medical care utilisation in South Carolina: A study protocol [J]. #N/A. 2019;9(7):e027688.

[5] Carra G, Salluh JIF, da Silva Ramos FJ, et al. Data-driven ICU management: Using Big Data and algorithms to improve outcomes [J]. J Crit Care. 2020;60:300-4.

[6] Kumar NMS, Eswari T, Sampath P, et al. Predictive Methodology for Diabetic Data Analysis in Big Data In: Vijayakumar V, Neelanarayanan V, editors. Big Data, Cloud and Computing Challenges [M]. Procedia Computer Science. 502015. p. 203-8.

[7] Razavian N, Blecker S, Schmidt AM, et al. Population-Level Prediction of Type 2 Diabetes From Claims Data and Analysis of Risk Factors [J]. #N/A. 2015;3(4):277-87.

[8] Fitzgerald RC. Big data is crucial to the early detection of cancer [J]. Nat Med. 2020;26(1):19-20.

[9] Fridley BL, Koestler DC, Godwin AK. Individualizing care for ovarian cancer patients using big data [J]. J Natl Cancer Inst. 2014;106(5).
[10] Cammarota G, Ianiro G, Ahern A, et al. Gut microbiome, big data and machine learning to promote precision medicine for cancer [J]. Nat Rev Gastroenterol Hepatol. 2020;17(10):635-48.
[11] Harmon SA, Sanford TH, Xu S, et al. Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets [J]. Nat Commun. 2020;11(1):4080.
[12] Mei X, Lee HC, Diao KY, et al. Artificial intelligence-enabled rapid diagnosis of patients with COVID-19 [J]. Nat Med. 2020;26(8):1224-8.
[13] Wang CJ, Ng CY, Brook RH. Response to COVID-19 in Taiwan: Big Data Analytics, New Technology, and Proactive Testing [J]. Jama. 2020;323(14):1341-2.
[14] Jensen PB, Jensen LJ, Brunak S. Mining electronic health records: towards better research applications and clinical care [J]. Nature reviews Genetics. 2012;13(6):395-405.
[15] Andreu-Perez J, Poon CC, Merrifield RD, et al. Big data for health [J]. IEEE J Biomed Health Inform. 2015;19(4):1193-208.
[16] O'Connor S, Zhang M, Honey M, et al. Digital professionalism on social media: A narrative review of the medical, nursing, and allied health education literature [J]. Int J Med Inf. 2021;153:104514.
[17] Attai DJ, Anderson PF, Fisch MJ, et al. Risks and benefits of Twitter use by hematologists/oncologists in the era of digital medicine [J]. Seminars in hematology. 2017;54(4):198-204.
[18] Epstein RM, Hundert EM. Defining and assessing professional competence [J]. Jama. 2002;287(2):226-35.
[19] Cain J, Romanelli F. E-professionalism: a new paradigm for a digital age [J]. Currents in Pharmacy Teaching and Learning. 2009;1(2):66-70.
[20] Hao X, Peng X, Ding X, et al. Application of digital education in undergraduate nursing and medical interns during the COVID-19 pandemic: A systematic review [J]. Nurse Educ Today. 2022;108:105183.
[21] Kailey BS, Dev D, Sado DM, et al. CardioWebinar: the evolution of digital education during the COVID-19 pandemic [J]. Heart. 2021;107(24):2004-5.
[22] Cruess R, McIlroy JH, Cruess S, et al. The Professionalism Mini-Evaluation Exercise: A Preliminary Investigation [J]. 2006;81(10):S74-S8.
[23] Cruess R, McIlroy JH, Cruess S, et al. The Professionalism Mini-Evaluation Exercise: A Preliminary Investigation [J]. Academic Medicine. 2006;81(10).
[24] Norcini JJ, Blank LL, Arnold GK, et al. The mini-CEX (clinical evaluation exercise): a preliminary investigation [J]. Ann Intern Med. 1995;123(10):795-9.
[25] Parthiban N, Boland F, Fadil Azim DH, et al. Asian medical students' attitudes towards professionalism [J]. Med Educ Online. 2021;26(1):1927466.
[26] Tsugawa Y, Ohbu S, Cruess R, et al. Introducing the Professionalism Mini-Evaluation Exercise (P-MEX) in Japan: results from a multicenter, cross-sectional study [J]. Acad Med. 2011;86(8):1026-31.
[27] Bajwa NM, Nendaz MR, Posfay-Barbe KM, et al. A Meaningful and Actionable Professionalism Assessment: Validity Evidence for the Professionalism Mini-Evaluation Exercise (P-MEX) Across 8 Years [J]. Acad Med. 2021;96(11s):S151-s7.

[28] Wu J, Wang J, Nicholas S, et al. Application of Big Data Technology for COVID-19 Prevention and Control in China: Lessons and Recommendations [J]. J Med Internet Res. 2020;22(10):e21980.

[29] Shen T, Wang C. Big Data Technology Applications and the Right to Health in China during the COVID-19 Pandemic [J]. International journal of environmental research and public health. 2021;18(14).

[30] Li X, Cong Y. A Systematic Literature Review of Ethical Challenges Related to Medical and Public Health Data Sharing in China [J]. Journal of Empirical Research on Human Research Ethics. 2021:15562646211040299.

[31] Liang W, Yao J, Chen A, et al. Early triage of critically ill COVID-19 patients using deep learning [J]. Nature communications. 2020;11(1):3543-.

[32] Meo SA, Sattar K, Ullah CH, et al. Progress and prospects of medical education research in Asian Countries [J]. Pakistan journal of medical sciences. 2019;35(6):1475-81.

[33] Schmid RD, Xiong X. Biotech in China 2021, at the beginning of the 14th five-year period ("145") [J]. Applied microbiology and biotechnology. 2021;105(10):3971-85.

[34] Wu H, Li W, Xie A, et al. Funding of health professional education: China's 20-year process and a global comparison [J]. Medical education. 2021.

[35] Silva Junior G, Askari M, Oliveira J. Digital Health and Possible Solutions to Improve the Care in the Field of Nephrology [J]. Contributions to nephrology. 2021;199:1-15.

[36] Doukas DJ, McCullough LB, Wear S. Perspective: Medical education in medical ethics and humanities as the foundation for developing medical professionalism [J]. Acad Med. 2012;87(3):334-41.

[37] Liu S, Chen Y, Li S, et al. What Are the Important Factors Influencing the Recruitment and Retention of Doctoral Students in a Public Health Setting? A Discrete Choice Experiment Survey in China [J]. International journal of environmental research and public health. 2021;18(18).

[38] Jingang A. Which future for doctors in China? [J]. Lancet. 2013;382(9896):936-7.

[39] Walkowiak D, Domaradzki J. Are rare diseases overlooked by medical education? Awareness of rare diseases among physicians in Poland: an explanatory study [J]. Orphanet journal of rare diseases. 2021;16(1):400.

[40] Domaradzki J, Walkowiak D. Knowledge and Attitudes of Future Healthcare Professionals Toward Rare Diseases [J]. Front Genet. 2021;12:639610.

Figure Legends

Figure 1. Basic knowledge for big data platform (A) before and (B) after the workshop.

* means that there is a significant difference before and after the workshop.

Figure 2. Students' attitude towards big data platform (A) before and (B) after the workshop.

* means that there is a significant difference before and after the workshop.

Figure 3. The professionalism for students (A) before and (B) after the workshop.

* means that there is a significant difference before and after the workshop.

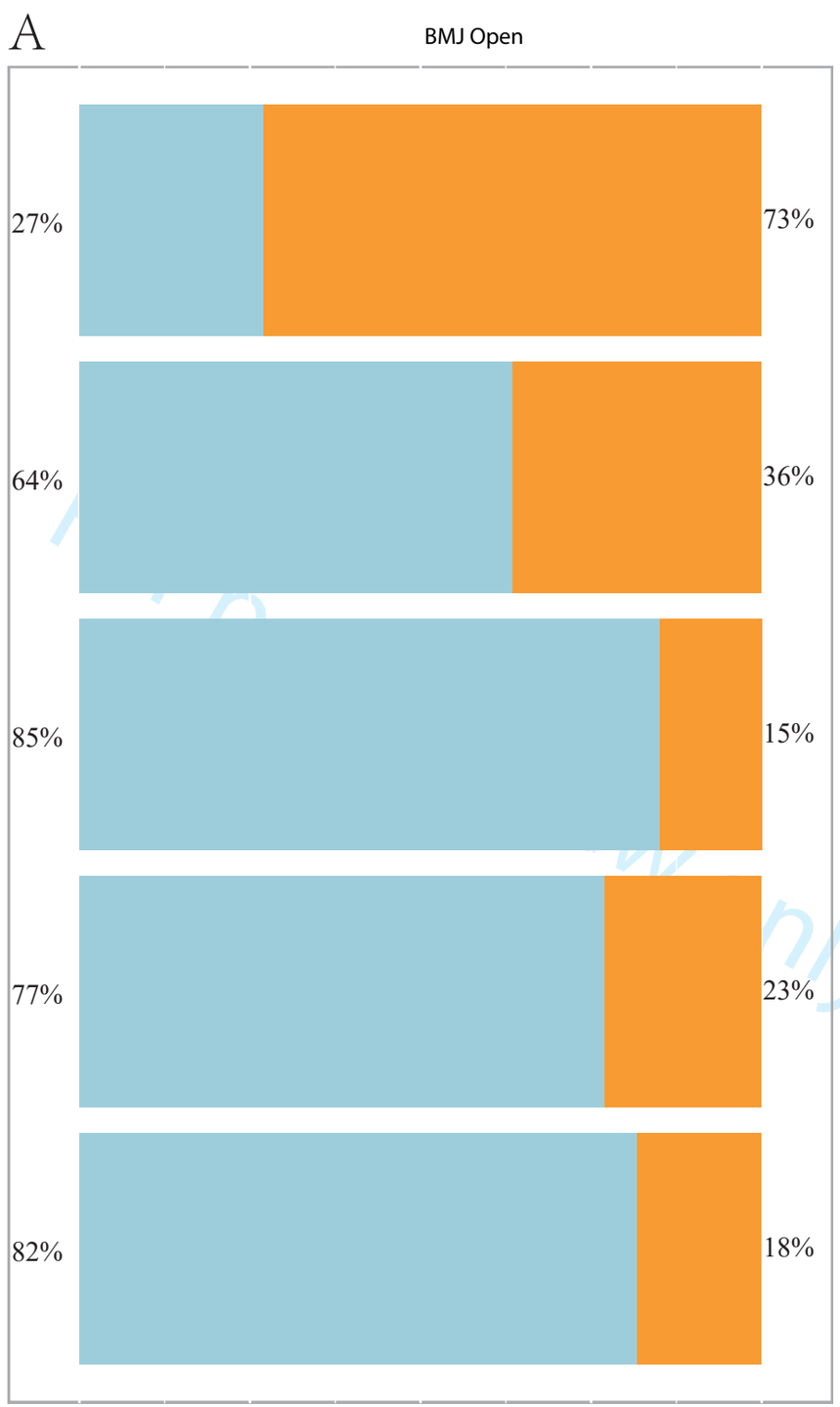
*I know the content of big data platform

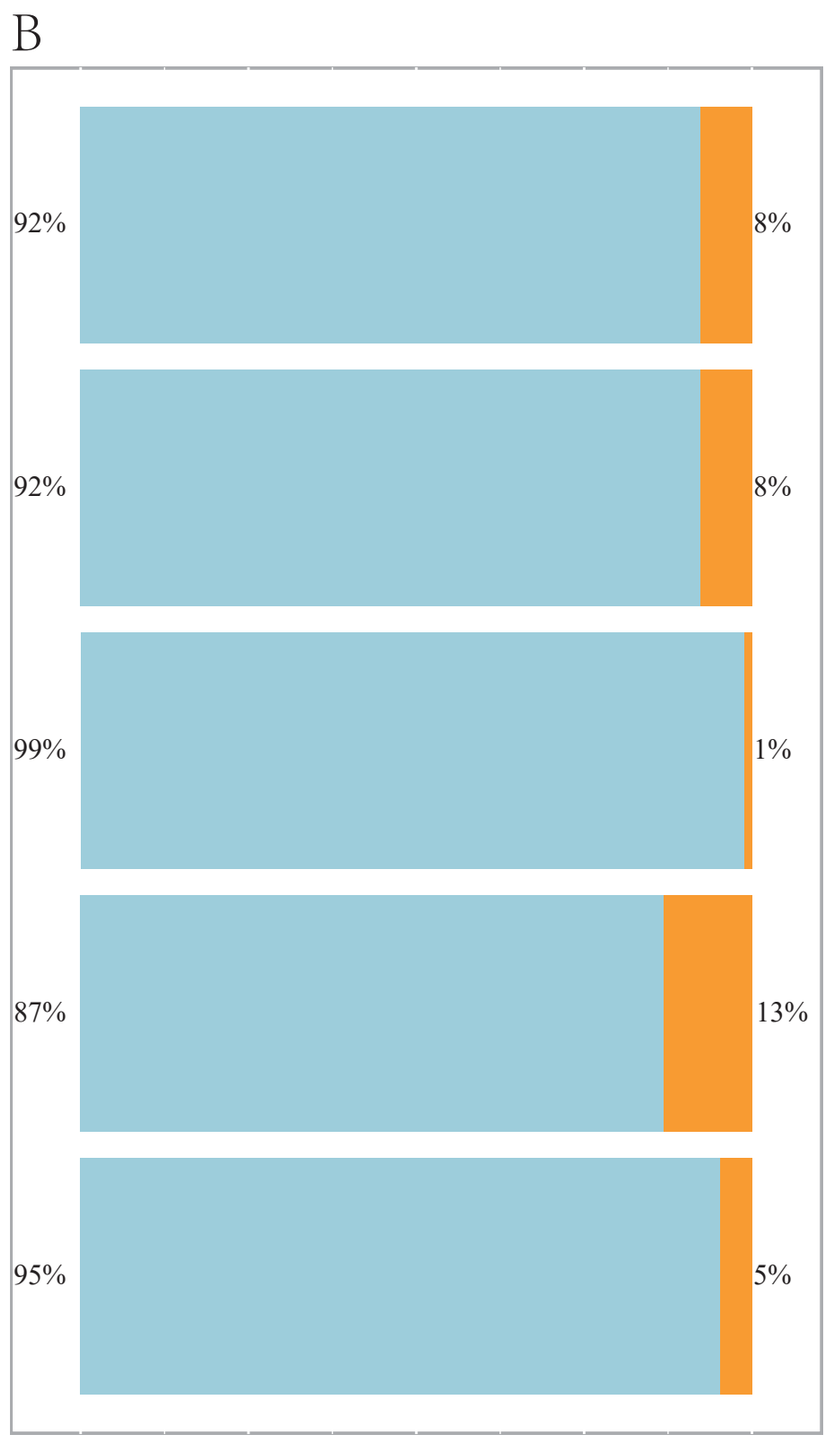
*I know the potential applications of big data technology

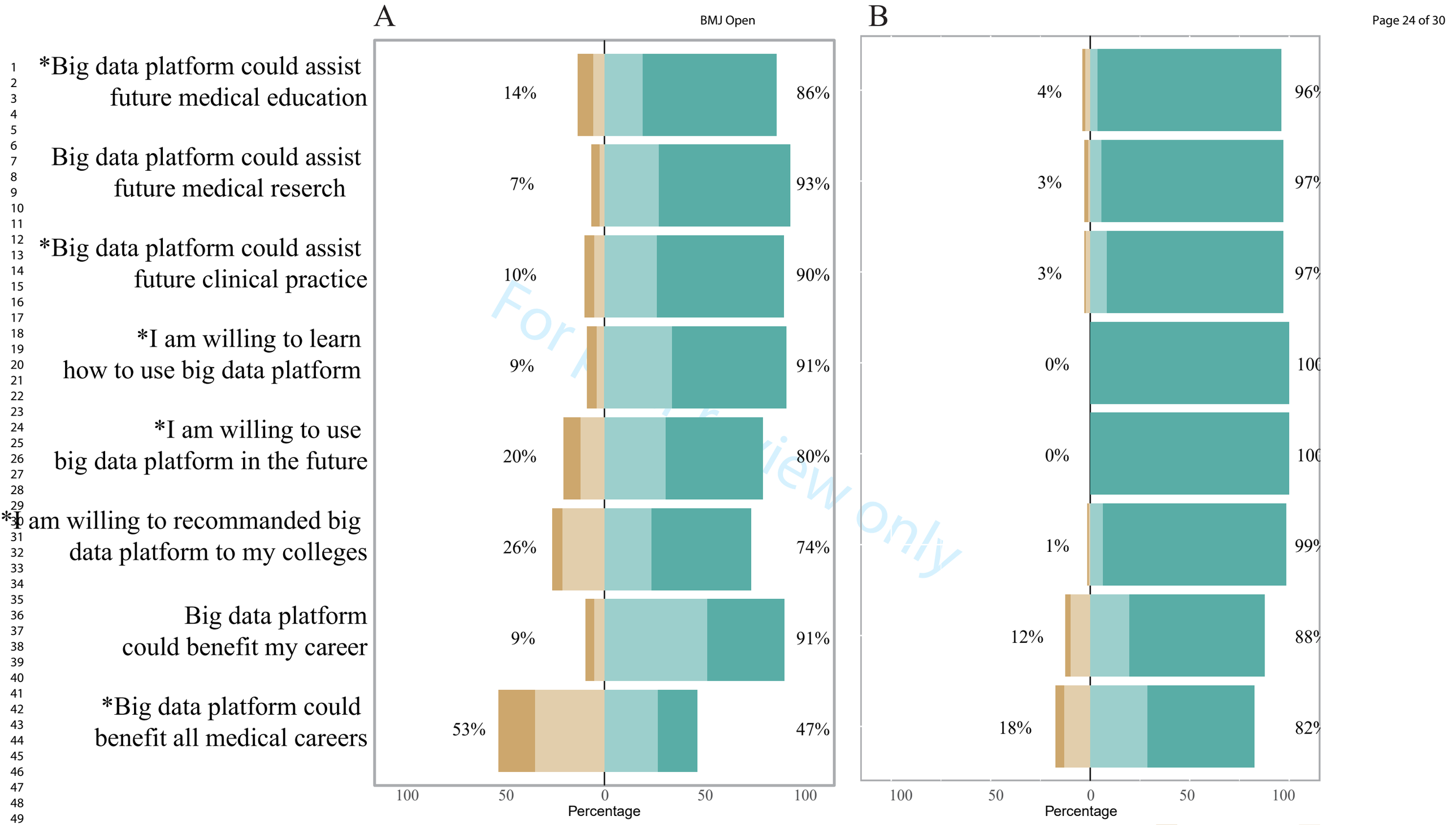
*Big data technology is practical in medical education

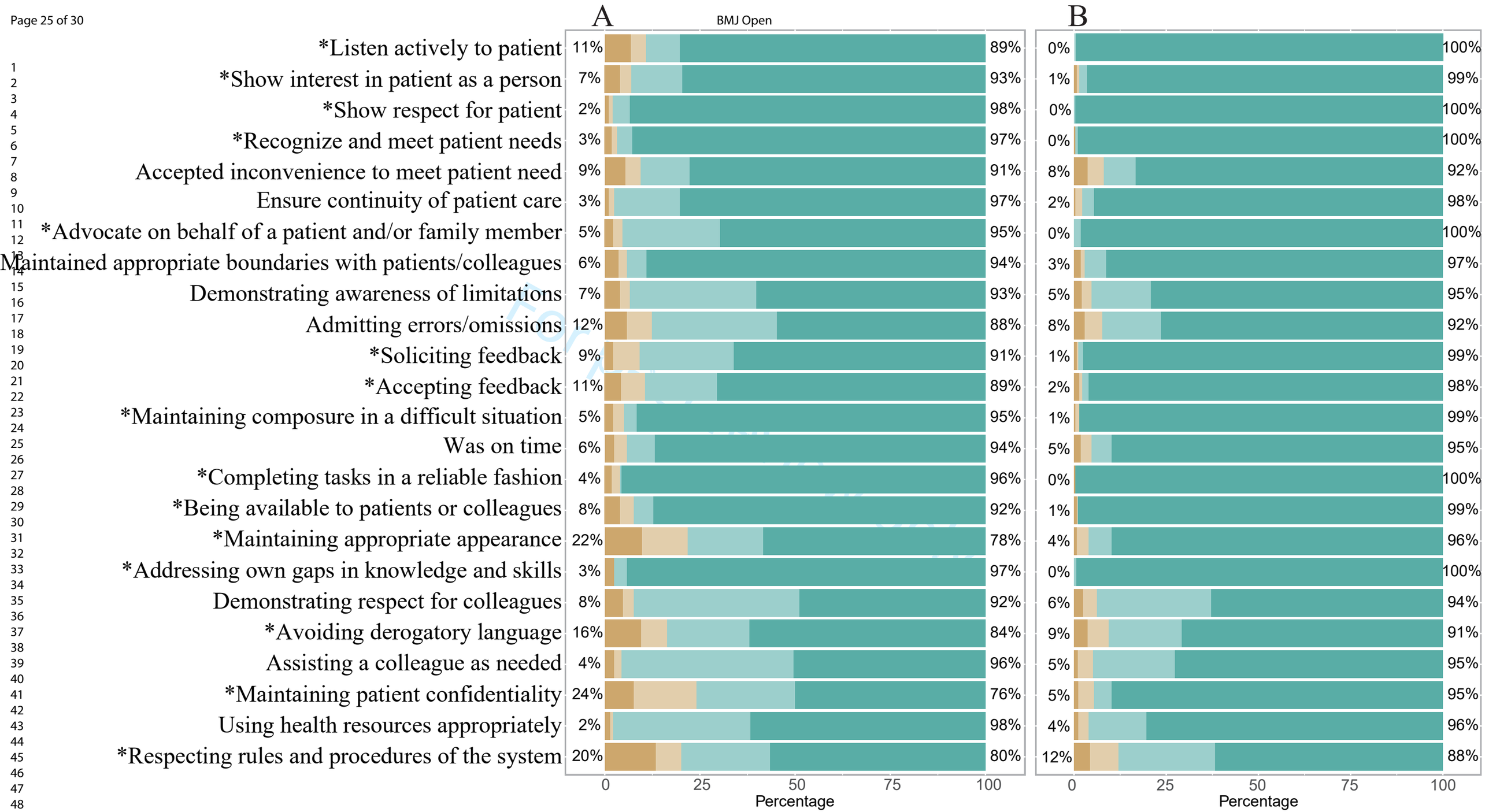
*Big data technology is practical in clinical practice

*Big data technology is practical in scientific research









Supplemental tables

Table S1, basic knowledge for big data platform before and after the workshop.

	Before Workshop		After Workshop		
Items	YES, n	NO, n	YES, n	NO, n	*P*
I know the content of big data platform	74	200	253	21	*<0.0001*
I know the potential applications of big data technology	174	100	253	21	*<0.0001*
Big data technology is practical in medical education	233	41	271	3	*<0.0001*
Big data technology is practical in clinical practice	211	63	238	36	*0.0039*
Big data technology is practical in scientific research	224	50	261	13	*<0.0001*

Table S2, students' attitude towards big data platform before and after the workshop.

	Before Workshop						After Workshop						*P*
Items	Totally disagree, n	Disagree, n	Agree, n	Totally agree, n	Mean	SD	Totally disagree, n	Disagree, n	Agree, n	Totally agree, n	Mean	SD	
Big data platform could assist future medical education	21	16	53	184	3.50	0.910	4	7	10	253	3.87	0.504	*0.0002*
Big data platform could assist future scientific research	11	7	75	181	3.55	0.735	5	3	16	250	3.86	0.499	*0.0705*
Big data platform could assist future clinical practice	13	14	72	175	3.49	0.799	2	6	23	243	3.85	0.464	*0.0017*
I am willing to learn how to use big data platform	13	11	93	157	3.43	0.783	0	0	0	274	4	0	*<0.0001*

I am willing to use big data platform in the future	23	33	84	134	3.20	0.953	0	0	0	274	4	0	*<0.0001*
I am willing to recommended big data platform to my colleagues	14	58	65	137	3.19	0.941	2	2	18	252	3.90	0.388	*<0.0001*
Big data platform could benefit my career	12	14	142	106	3.25	0.744	7	27	54	186	3.53	0.776	*0.3382*
Big data platform could benefit all medical careers	50	96	74	54	2.48	1.006	12	36	79	147	3.32	0.863	*<0.0001*

SD, Standard Deviation.

Table S3, the professionalism for students before and after the workshop.

	Before workshop				After workshop				*P*
Items	Not at all important, n	Not important, n	Important, n	Very important, n	Not at all important, n	Not important, n	Important, n	Very important, n	
Doctor-patient relationship skills									
Listen actively to patient	19	11	24	220	0	0	1	273	*<0.0001*
Show interest in patient as a person	11	8	37	218	2	2	6	264	*0.0029*
Show respect for patient	3	3	12	256	0	0	1	273	*0.0401*
Recognize and meet patient needs	5	4	11	254	1	0	2	271	*0.0255*
Accepted inconvenience to meet patient need	15	11	35	213	10	12	24	228	*0.6503*
Ensure continuity of patient care	3	4	47	220	1	5	9	259	*1*

Advocate on behalf of a patient and/or family member	6	7	70	191	0	0	5	269	*0.0008*
Maintained appropriate boundaries with patients/colleagues	10	6	14	244	5	3	16	250	*0.1440*
Reflective skills									
Demonstrating awareness of limitations	11	7	91	165	6	7	44	217	*0.4595*
Admitting errors/omissions	16	18	90	150	8	13	44	209	*0.0880*
Soliciting feedback	6	19	68	181	2	1	4	267	*<0.0001*
Accepting feedback	12	17	52	193	4	2	5	263	*0.0001*
Maintaining composure in a difficult situation	6	8	9	251	1	3	0	270	*0.0310*
Time management									
Was on time	7	9	20	238	5	8	15	246	*0.7027*
Completing tasks in a reliable fashion	5	6	1	262	1	0	0	273	*0.0086*
Being available to patients or colleagues	11	10	14	239	2	1	0	271	*0.0004*
Interprofessional relationship skills									
Maintaining appropriate appearance	27	33	54	160	2	9	17	246	*<0.0001*
Addressing own gaps in knowledge and skills	7	0	9	258	0	0	2	272	*0.0225*
Demonstrating respect for colleagues	13	8	119	134	7	10	85	172	*0.6139*
Avoiding derogatory language	26	19	59	170	10	16	54	194	*0.0220*
Assisting a colleague as needed	7	5	124	138	3	11	61	199	*0.8407*

Maintaining patient confidentiality	21	45	71	137	3	12	13	246	*<0.0001*
Using health resources appropriately	4	2	99	169	3	8	43	220	*0.3244*
Respecting rules and procedures of the system	37	18	64	155	12	21	72	169	*0.0146*

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4
Objectives	3	State specific objectives, including any prespecified hypotheses	4-5
Methods			
Study design	4	Present key elements of study design early in the paper	5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	5
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case	NA
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	6-7
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	6-7
Bias	9	Describe any efforts to address potential sources of bias	NA
Study size	10	Explain how the study size was arrived at	6
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	7-8
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	7-8
		(*b*) Describe any methods used to examine subgroups and interactions	NA
		(*c*) Explain how missing data were addressed	7-8
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	7-8
		(*e*) Describe any sensitivity analyses	NA

Continued on next page

Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	8
		(b) Give reasons for non-participation at each stage	All participants were eligible and included in this study.
		(c) Consider use of a flow diagram	NA
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	8-9
		(b) Indicate number of participants with missing data for each variable of interest	Very high response rates with no missing data, as shown in supplementary data.
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	NA
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	NA
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	NA
		Cross-sectional study—Report numbers of outcome events or summary measures	8
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	In all tables and figures.
		(*b*) Report category boundaries when continuous variables were categorized	In all tables and figures.
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	NA
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	NA
Discussion			
Key results	18	Summarise key results with reference to study objectives	13
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	15
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	16
Generalisability	21	Discuss the generalisability (external validity) of the study results	15
Other information			

Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	16

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

Oncologic big data platform for promoting digital competencies and professionalism in Chinese medical students: a cross-section study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2022-061015.R3
Article Type:	Original research
Date Submitted by the Author:	24-Aug-2022
Complete List of Authors:	Liu, Jiahao; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Jiao, Xiaofei; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Zeng, Shaoqing; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Li, Huayi; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Jin, Ping; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Chi, Jianhua; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Liu, Xingyu; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Yu, Yang; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Ma, Guanchen; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Zhao, Yingjun; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Li, Ming; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Peng, Zikun; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Huo, Yabing; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Gao, Qing-Lei ; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology
Primary Subject Heading:	Medical education and training
Secondary Subject Heading:	Medical education and training
Keywords:	MEDICAL EDUCATION & TRAINING, Gynaecological oncology < GYNAECOLOGY, Gynaecological oncology < ONCOLOGY, Health informatics < BIOTECHNOLOGY & BIOINFORMATICS

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Oncologic big data platform for promoting digital competencies and professionalism in Chinese medical students: a cross-section study

Jiahao Liu[1,2#], Xiaofei Jiao[1,2#], Shaoqing Zeng[1,2], Huayi Li[1,2], Ping Jin[1,2], Jianhua Chi[1,2], Xingyu Liu[1,2], Yang Yu[1,2], Guanchen Ma[1,2], Yingjun Zhao[1,2], Ming Li[1,2], Zikun Peng[1,2], Yabing Huo[1,2], Qinglei Gao[1,2*]

[1] Department of Obstetrics and Gynecology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, Hubei, P. R. China.
[2] Cancer Biology Research Center (Key Laboratory of the Ministry of Education), Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, Hubei, P. R. China.

[#] These authors contribute equally.

* Correspondence to:
Prof. Qinglei Gao, Cancer Biology Research Center (Key Laboratory of the Ministry of Education), Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, 1095 Jiefang Ave, Wuhan 430030, Hubei, P. R. China. Tel: +86-27-83662681; Fax: +86-27-83662681; E-mail: qingleigao@hotmail.com.

Word count (excluding title page, abstract, references, figures and tables): 3,202

Abstract

Objectives: Advancements in big data technology are reshaping the health care system in China. This study aims to explore the role of medical big data in promoting digital competencies and professionalism among Chinese medical students.

Design, setting and participants: This study was conducted among 274 medical students who attended a workshop on medical big data conducted on July 8, 2021 in Tongji Hospital. The workshop was based on the first nationwide multifunction gynecologic oncology medical big data platform in China, at the National Union of Real-World Gynecologic Oncology Research & Patient Management Platform (NUWA platform).

Outcome measures: Data on knowledge, attitudes toward big data technology and professionalism were collected before and after the workshop. We have measured the four skill categories: doctor–patient relationship skills, reflective skills, time management, and interprofessional relationship skills using the Professionalism Mini–Evaluation Exercise (P-MEX) as a reflection for professionalism.

Results: A total of 274 students participated in this workshop and completed all the surveys. Before the workshop, only 27% of them knew the detailed content of medical big data platforms, and 64% knew the potential application of medical big data. The majority of the students believed that big data technology is practical in their clinical practice (77%), medical education (85%), and scientific research (82%). Over 80% of the participants showed positive attitudes toward big data platforms. They also exhibited sufficient professionalism before the workshop. Meanwhile, the workshop significantly promoted students' knowledge of medical big data ($P < 0.05$), and led to more positive attitudes toward big data platforms and higher levels of professionalism.

Conclusions: Chinese medical students have primitive acquaintance and positive attitudes toward big data technologies. The NUWA platform-based workshop may potentially promote their understanding of big data and enhance professionalism, according to the self-measured P-MEX scale.

Keywords: Big data, Chinese medical education, Digital competencies,

Professionalism, Workshop

Strengths and limitations of this study

- This study depicted the knowledge and professionalism of medical big data among Chinese medical students in the era of big data for the first time.
- This study provided a vivid example of how big data can assist medical education, based on the first nationwide gynecologic oncology medical big data platform in China (NUWA platform).
- The generalization of the findings is limited by the lack of sample representation and the short follow-up time.

Background

Electronification of medical records is the signature of the modern health care system.[1-3] Massive clinical and omics data have been produced to enable more detailed depictions of patients and diseases. Medical big data is thus reshaping our appreciation of the modern medical system. Technological advancements in data storage, processing, and analysis accelerated the clinical application of big data-driven products and contributed to personalized disease management,[4, 5] early diagnosis,[6-8] and treatment decision.[9, 10] Especially in the COVID-19 pandemic setting, achieving rapid application of medical big data would meet the pressing clinical need to predict the progression of diseases using data characteristics.[11-13]

Big data has brought new challenges for doctors.[14] Information overload is a challenge for every healthcare worker, since they have to acclimate to the nature of big data, including extraordinary value, volume, velocity, variety, and variability.[15] Challenges can also be expected regarding medical professionalism in the age of big data. Because big data values realistic health-related information more than ever, the illusion that digital data outweigh face-to-face physician–patient interactions may motivate doctors to ignore the importance of professionalism.[16, 17] There is no precise definition of medical professionalism, which is reflected in the attitudes and behaviors directly related to clinical practice. Epstein RM et al. proposed the definition of professional competence as wisely using communication, knowledge, technical skills, clinical reasoning, emotions, values, and reflection in daily practice.[18] And Cain J et al. summarized professionalism as a series of attributes, such as altruism, respect, honesty, and so on.[19] Some studies have shown that digital medical education based on big data plays a positive effect in promoting professionalism.[16, 20, 21] However, there is a lack of research on the attitude and professionalism of Chinese healthcare workers regarding medical big data. Therefore, extra lectures and workshops about obtaining insights into big data and remaining respectful to patients are necessary.

In this study, we aimed to depict the knowledge of medical big data and professionalism in the era of big data among Chinese medical students. What's more,

we conducted a workshop for Chinese clinical and preclinical students to improve their knowledge of medical big data and their professionalism in the big data era, based on the National Union of Real-World Gynecologic Oncology Research & Patient Management Platform (NUWA platform), which is the first nationwide multifunction gynecologic oncology medical big data platform in China.

Methods

Study design

A workshop on the introduction of medical big data was conducted among clinical and preclinical students in Tongji Medical School. We conducted this study on Chinese medical students' learning and the application of big data in healthcare using pre- and post-course questionnaires.

Workshop design

The workshop was conducted on July 8, 2021 in Tongji Hospital, Tongji Medical School, Huazhong University of Science and Technology, Wuhan, China. It was part of the series class of "Medical big data platform learning and applying".

The workshop was delivered by two senior doctors (QLG and SQZ), who have over five years of experience in medical big data platform development and application. The workshop consisted of three major parts, (i) a lecture on the content and application of medical big data; (ii) a lecture on professionalism for doctors in the big data era; and (iii) learners being allowed to explore the NUWA platform freely according to their interest. The first and second sections lasted for approximately 45 minutes, and the third section lasted for 30 minutes.

In section one, five major parts were included in lectures given by the senior doctors, (i) the development of medical big data in China and the world, (ii) the application and potential of big data in clinical practice, medical education, and scientific research, (iii) the construction and content of the NUWA platform, (iv) how to use the NUWA

platform, and (v) plans for the NUWA platform development. For the second section, the following four items were discussed: (i) the attitude health care workers should have when communicating with patients and colleagues, considering that big data have already reshaped our medical system, (ii) the importance of detailed and continuous patient information for the development of big data platforms, (iii) how to protect patient privacy on the databases, and (iv) how to use the NUWA platform to satisfy patients' needs. For Section 3, all students were given a temporary account for the NUWA platform and had access to all deidentified patient information. They were allowed to explore the medical data freely and view structured health care information for half an hour.

Data collection

Clinical and preclinical students in Tongji Medical School were invited to attend a 2-hour class by email or roadshows between June 8, 2021, and July 7, 2021. The contents and speakers for the workshop were presented in the email or during the roadshow. In addition, all 5th- (n=50, preclinical) and 6th-grade (n=49, clinical) students of 8-year undergraduate education in Tongji Medical School took part in this workshop as an additional course.

The participants completed two surveys, one before the workshop and the other after the workshop. Both questionnaires contain their basic knowledge and attitude toward big data technology and a Professionalism Mini-Evaluation Exercise (P-MEX) instrument. The P-MEX is developed from the mini-Clinical Examination Exercise (mini-CEX) format by Cruess R et al. in 2006 to evaluate professionalism in clinical training.[22] It consists of 24 items representing four skill categories: doctor–patient relationship skills, reflective skills, time management, and interprofessional relationship skills.[23, 24] And the reliability and validity of P-MEX have been confirmed in both Eastern and Western cultural backgrounds.[25-27] The attitude scale is a self-created scale focused on measuring participants' pedagogic evaluation, acceptance and expectation of the big data platform. It consisted of 8 questions: (i) Big data platform

could assist future medical education, (ii) Big data platform could assist future medical research, (iii)Big data platform could assist future clinical practice, (iv) I am willing to learn how to use big data platform, (v) I am willing to use big data platform in the future, (vi) I am willing to recommend big data platform to my colleagues, (vii) Big data platform could benefit my career, and (viii) Big data platform could benefit all medical careers. For each question, students chose from "totally agree", "agree", "disagree", and "totally disagree". For the applicability of conclusions, "totally agree" and "agree" were regarded as "positive attitudes", and "disagree" and "totally disagree" were considered as "negative attitudes". In addition, the first survey also contained students' baseline information, and the second survey contained a satisfaction questionnaire. Meanwhile, the browsing histories of participants in Section 3 were also recorded and analyzed to reflect students' interests.

The data collection and analysis plan were acknowledged and agreed upon by all participants at the beginning of the workshop, and the study was approved by the Research Ethics Commission of Tongji Hospital of Huazhong University of Science and Technology (2020-S201).

NUWA platform

The NUWA platform is the first nationwide Gynecological Oncology data-sharing platform launched by the National Clinical Research Center for Gynecological Oncology in August 2019. This platform integrated inpatient/outpatient clinical data, genomic data, and follow-up data to develop a patient-level longitudinal clinico-genomic database. Information was deidentified and extracted from electronic medical records. A rigorous data quality check was performed to ensure the accuracy of the data entries. Since its foundation in 2019, 17 first-class hospitals from different provinces or cities in China have participated in the NUWA platform until August 2021.

Statistical analysis

Descriptive statistics were presented by counts and percentages to describe the

demographic information. The Chi-square test was used to compare the changes in knowledge and understanding of big data and professionalism before and after the workshop. To acquire more practical results, the categories "not at all important" and "not important" were combined for the analysis, as well as "important" and "very important". *P values <0.05* were considered statistically significant. The data were analyzed using R Version 4.03.

Patient and Public Involvement

No patients or members of the public were involved in this study.

Results

Characteristics of students

A total of 274 students participated in this workshop and completed two surveys. All of them were included in the final analysis. Participants were aged between 22 and 28, with 148 (54.0%) preclinical students and 126 (46.0%) clinical students. Among them, 130 (47.4%) were males and 144 (52.4%) were females (**Table 1**).

The majority of participants (207, 75.5%) knew of at least one big data platform in China or the world. Over two-thirds of them (183, 66.8%) also acknowledged its application. However, only 17.2% (47) of them had been involved in any project related to medical big data (**Table 1**).

Table 1. Baseline characteristics for participants.

Characteristics	Number of participants, n (%)
Age, years	
22	27 (9.9%)
23	16 (5.8%)
24	67 (24.5%)
25	66 (24.1%)
26	50 (18.2%)

	27	28 (10.2%)
	28	20 (7.3%)
Gender		
	male	130 (47.4%)
	female	144 (52.6%)
Study stage		
	pre-clinical	148 (54.0%)
	clinical	126 (46.0%)
Acknowledgment of any kind of big data platform		
	yes	207 (75.5%)
	no	67 (24.5%)
Know the applications of big data technology		
	yes	183 (66.8%)
	no	91 (33.2%)
Involved in any big data-related projects		
	yes	47 (17.2%)
	no	227 (82.8%)
Total		274 (100%)

Knowledge of big data platforms

Before the workshop, approximately a quarter of the students (74, 27%) knew the detailed content of at least one medical big data platform, while nearly 64% (174) knew the potential application of medical big data platforms. After attending the lessons, almost all students could understand the content (253, 92%) and the potential application (253, 92%) of medical big data (**Figure 1 and Table S1**).

Regardless of the survey completed (pre- or post-course), the majority of students

believed that big data technology is practical in medical education, clinical practice, and scientific research (85%, 77%, 82% before the workshop; 99%, 87%, 95% after the workshop, respectively) (**Figure 1 and Table S1**).

Students' attitudes toward big data platforms

Even before the workshop, most students hold a positive attitude on the potential of the big data platform in promoting medical education (237, 86%), medical research (256, 93%), and clinical practices (247, 90%). They were also willing to learn about and use the big data platform (250, 91%, and 218, 80%, respectively). Of them, 74% (202) were enthusiastic about introducing big data platforms to their colleagues. Interestingly, most students were convinced that the big data platform could benefit their careers (248, 91%), but they were not sure if it would yield the same effect on the others (128, 47%) (**Figure 2 and Table S2**).

When the workshop was completed, almost all students had positive attitudes toward the big data platform (**Figure 2 and Table S2**). However, there were still 12% (34) and 18% (48) of students who were not sure if the big data platform could benefit their and others' medical careers, respectively.

Professionalism

Generally, positive attitudes toward all the professionalism items were demonstrated, with more than three-quarters of the students agreeing that all professionalism attributes were "important" or "very important" before the workshop. The three items that most students thought were "not at all important" or "not important" were maintaining patient confidentiality (66, 24% of students chose "not important" or "not at all important"), maintaining appropriate appearance (60, 22% students chose "not important" or "not at all important") and respecting rules and procedures of the system (55, 20% students chose "not important" or "not at all important") (**Figure 3A and Table S3**).

After the lectures, students exhibited a more positive attitude on most

professionalism items ($p<0.05$). The most significant improvements occurred for "Maintaining patient confidentiality" (from 76% to 95%), "Listen actively to patient" (from 89% to 100%), and "Accepting feedback" (from 89% to 98%) (**Figure 3B and Table S3**).

Interest in medical data

In the free exploration section, most students (253, 92.3%) viewed the patients' hospitalization logs. A total of 76.3% (209) of students were interested in the patients' history of illness. Meanwhile, many students were interested in medicine usages (56.2%), surgery reports (52.9%), and diseases of rare pathologies (58.6%). Only approximately one-third of students viewed image diagnosis (37.6%) and follow-up records (32.5%) (**Table 2**).

Table 2. Browser records in the free-exploration section.

Content	Number of participants, n (%)
History of Illness	
yes	209 (76.3%)
no	65 (23.7%)
Hospitalization logs	
yes	253 (92.3%)
no	21 (7.7%)
Medicine usage	
yes	154 (56.2%)
no	120 (43.8%)
Surgery record	
yes	145 (52.9%)
no	129 (47.1%)
Image diagnosis	
yes	103 (37.6%)

no	171 (62.4%)
Follow-up records	
yes	89 (32.5%)
no	185 (67.5%)
Rare pathologies	
yes	160 (58.4%)
no	114 (41.6%)
Total	274 (100%)

Satisfactory survey

Upon completing the survey, 95.6% (262) of participants were "satisfied" or "extremely satisfied" with this workshop, and only 7 out of 274 students were not at all satisfied. However, we failed to obtain feedback from them in the next three months. Most students thought the workshop was informative (249, 90.9%) and understandable (255, 94.1%). The majority of students were also willing to recommend this seminar to other students (257, 93.8%) and to participate in similar classes in the future (254, 92.7%). Regarding the duration of this workshop, only 7 (2.6%) students thought that 2 hours was too long (**Table 3**).

Table 3. Participants 'answers to the workshop satisfaction survey.

	Not at all, n (%)	No, n (%)	Yes, n (%)	Yes, extremely, n (%)	Mean	SD
Overall, are you satisfied with this course	7 (2.6%)	5 (1.8%)	57 (20.8%)	205 (74.8%)	3.68	0.640
Did you think the course are informative?	10 (3.6%)	15 (5.5%)	101 (36.9%)	148 (54%)	3.41	0.757
Did you think the duration of this courses is too long	121 (44.2%)	118 (43.1%)	28 (10.2%)	7 (2.6%)	2.07	0.954

Was the course understandable for you	11 (4%)	8 (2.9%)	121 (44.2%)	134 (48.9%)	3.38	0.733
Would you recommend these courses to other students?	6 (2.2%)	11 (4%)	144 (52.6%)	113 (41.2%)	3.33	0.659
Are you willing to take part in similar courses in the future?	6 (2.2%)	14 (5.1%)	35 (12.8%)	219 (79.9%)	3.70	0.666

SD, Standard Deviation.

Discussion

Our study demonstrated that Chinese medical students have little knowledge of and positive attitudes toward big data technologies. They also yield expertise in professionalism. Furthermore, workshops based on big data platforms could further strengthen their digital competencies and improve doctor–patient communication capabilities, which would lead to better fitness during the expansion of medical big data.

In this workshop, the basic knowledge of big data platforms in China was measured. Although a large fraction of students knew about the big data platform and its utility, only a few of them could apprehend its composition or how to use it. This situation may be caused by the recent rapid development of big data technology in China and comparatively lagging relevant education.[28, 29] In addition, ethical challenges that hinder medical and public health data sharing may have also exacerbated the situation.[30] Therefore, aside from conducting big data-related workshops, promoting medical and public health data sharing policies and evolving relevant legal and ethical implications are also of great importance.

Students' enthusiasm toward medical big data was also taken into consideration. As expected and consistent with previous reports,[25] the majority of students exhibited positive attitudes toward the big data platform and were willing to acquire further

proficiency. They believed that big data would play a vital role in future medical education, clinical practice, and scientific research, which are the main tasks for all Chinese doctors. However, not all of them believed that big data could benefit all medical careers. The popularity of information technology in China recently may contribute to the big data-friendly intention of medical students.[31] Meanwhile, their limited understanding of related fields may hamper the belief that it could be used productively.[32-34] In the meantime, we were delighted to see that almost all students recognized the significance of big data after our workshop. We thus believe that proper education may be an efficient way to diminish misunderstandings and achieve full potential in big data platforms.

Professionalism was related to the development of big data in our study for the first time in China. Considering that big data can reshape medical activities in all aspects, including doctors' attitudes toward their colleagues and patients,[35] a lecture about "professionalism in the age of big data" was presented in the current workshop. Before the seminar, students demonstrated sufficient professionalism, similar to the findings of previous reports, with high scores on the majority of these items. Nevertheless, it is worth noting that nearly 1/4 of the participants did not pay enough attention to maintaining patient confidentiality. Protecting patients' privacy is an essential embodiment of medical ethics and humanities.[36] Participants mostly answered these questions from the perspective of big data users before our workshop, ignoring that those data represented thousands of actual patients. After we emphasized the importance of data privacy in the workshop, students realized that respect for patients is the foundation for improving their medical professionalism. Meanwhile, to enhance patient-data privacy protection, the privacy information, such as ID number, was privatized and converted into an alternative ID number using a hashing algorithm before being uploaded to the NUWA platform. Therefore, the personal information of all included patients in the NUWA platform is confidential and unreachable to all users. Several important professionalism levels were also significantly increased in many other elements after the workshop, indicating that this workshop could be a preliminary

attempt to promote professionalism when the prominence of digital data has changed the way we communicate and when doctors have to spend more time with electronic records than with patients. These changes our workshop brought deserve more attention in China since Chinese doctors are well-known to be overburdened.[37, 38] These improvements could also help increase the reliability of medical records and produce convincing and effective medical information.

Another interesting fact in the results is that approximately 60% of students noticed rare diseases in the free-exploration section, which is hardly involved in routine medical classes.[39, 40] We believe this is another strength of big data-based medical education. In traditional medical classes, it is arduous for teachers to grant detailed depictions for every type of rare disease. As a result, students may not have easy access to these exceptional cases and therefore are often not able to diagnose rare diseases when reviewing patients' medical records. The big data platform makes it possible for every user to conveniently browse through cases with rare pathologies, which would benefit patients as well as reduce misdiagnoses.

Our workshop had some limitations. The most important one is that the NUWA platform is still under development, and the omics data are not currently included. Therefore, another investigation should be conducted when the construction of NUWA has been completed. Meanwhile, we did not calculate the sample size before the conduction of this study, which may lead to potential bias. And the limited sample size may be another reason that restricts the popularization of our conclusion. Furthermore, narrow geographical distributions of study participants interfered with the generalization of results to other populations. Meanwhile, there are concerns that students who volunteered to participate in the workshop may be more enthusiastic about big data technology than those who did not. We thus included all 5th- (preclinical) and 6th-grade (clinical) students of 8-year undergraduate education in Tongji Medical School to make the study sample more representative. As for the measurement of attitudes and professionalism, a qualitative approach would have been more appropriate and should be considered in the future. Meanwhile, this workshop is too short to

produce fundamental improvements in students' attitudes and professionalism. In the future, more long-term studies are wanted to draw a clear conclusion.

Conclusion

This study depicted Chinese students' knowledge of medical big data for the first time and the NUWA platform-based workshop had the potential to improve their understanding of big data and enhance professionalism.

List of abbreviations

P-MEX, Professionalism Mini-Evaluation Exercise instrument. NUWA platform, the National Union of Real-World Gynecologic Oncology Research & Patient Management Platform.

Contributorship

QLG, JHL and XFJ designed the study and survey. QLG, JHL, XFJ, ML, ZKP, YBH and SQZ conducted the workshop. PJ, HYL, JHC, XYL, YY, and SQZ analyzed and interpreted the data. YJZ and GCM performed interpretation of data and discussion of findings. QLG conceptualized and designed the study, supervised the project, analyzed and interpreted the data, and wrote the paper. All authors approved the final version of the manuscript and agreed to the submission of this manuscript.

Funding statement

This work was supported by the Teaching Reform Project for Postgraduates of Huazhong University of Science and Technology (YY202006), the Major Project of Medical Education (20A0769), and Teaching Research Project of The Second Clinical School, Tongji Medical College, Huazhong University of Science and Technology (2021022).

Competing of interests

The authors declare that they have no competing interests.

Ethics approval

The data collection and analysis plan were acknowledged and agreed by all participants and the study has been approved by the Research Ethics Commission of Tongji Hospital of Huazhong University of Science and Technology (2020-S201).

Data sharing

The data used in the current study are available from the corresponding author on reasonable request via e-mail (qingleigao@hotmail.com).

Consent for publication

Not applicable.

Acknowledgements

Not applicable.

Reference

[1] Bezemer T, de Groot MC, Blasse E, et al. A Human(e) Factor in Clinical Decision Support Systems [J]. J Med Internet Res. 2019;21(3):e11732-e.

[2] Sim I. Two Ways of Knowing: Big Data and Evidence-Based Medicine [J]. Ann Intern Med. 2016;164(8):562-3.

[3] Wu WT, Li YJ, Feng AZ, et al. Data mining in clinical big data: the frequently used databases, steps, and methodological models [J]. Military Medical Research. 2021;8(1):44.

[4] Olatosi B, Zhang J, Weissman S, et al. Using big data analytics to improve HIV medical care utilisation in South Carolina: A study protocol [J]. #N/A. 2019;9(7):e027688.

[5] Carra G, Salluh JIF, da Silva Ramos FJ, et al. Data-driven ICU management: Using Big Data and algorithms to improve outcomes [J]. J Crit Care. 2020;60:300-4.

[6] Kumar NMS, Eswari T, Sampath P, et al. Predictive Methodology for Diabetic Data Analysis in Big Data In: Vijayakumar V, Neelanarayanan V, editors. Big Data, Cloud and Computing Challenges [M]. Procedia Computer Science. 502015. p. 203-8.

[7] Razavian N, Blecker S, Schmidt AM, et al. Population-Level Prediction of Type 2 Diabetes From Claims Data and Analysis of Risk Factors [J]. #N/A. 2015;3(4):277-87.

[8] Fitzgerald RC. Big data is crucial to the early detection of cancer [J]. Nat Med. 2020;26(1):19-20.

[9] Fridley BL, Koestler DC, Godwin AK. Individualizing care for ovarian cancer patients using big data [J]. J Natl Cancer Inst. 2014;106(5).

[10] Cammarota G, Ianiro G, Ahern A, et al. Gut microbiome, big data and machine learning to promote precision medicine for cancer [J]. Nat Rev Gastroenterol Hepatol. 2020;17(10):635-48.

[11] Harmon SA, Sanford TH, Xu S, et al. Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets [J]. Nat Commun. 2020;11(1):4080.

[12] Mei X, Lee HC, Diao KY, et al. Artificial intelligence-enabled rapid diagnosis of patients with COVID-19 [J]. Nat Med. 2020;26(8):1224-8.

[13] Wang CJ, Ng CY, Brook RH. Response to COVID-19 in Taiwan: Big Data Analytics, New Technology, and Proactive Testing [J]. Jama. 2020;323(14):1341-2.

[14] Jensen PB, Jensen LJ, Brunak S. Mining electronic health records: towards better research applications and clinical care [J]. Nature reviews Genetics. 2012;13(6):395-405.

[15] Andreu-Perez J, Poon CC, Merrifield RD, et al. Big data for health [J]. IEEE J Biomed Health Inform. 2015;19(4):1193-208.

[16] O'Connor S, Zhang M, Honey M, et al. Digital professionalism on social media: A narrative review of the medical, nursing, and allied health education literature [J]. Int J Med Inf. 2021;153:104514.

[17] Attai DJ, Anderson PF, Fisch MJ, et al. Risks and benefits of Twitter use by hematologists/oncologists in the era of digital medicine [J]. Seminars in hematology. 2017;54(4):198-204.

[18] Epstein RM, Hundert EM. Defining and assessing professional competence [J]. Jama. 2002;287(2):226-35.

[19] Cain J, Romanelli F. E-professionalism: a new paradigm for a digital age [J]. Currents in Pharmacy Teaching and Learning. 2009;1(2):66-70.

[20] Hao X, Peng X, Ding X, et al. Application of digital education in undergraduate nursing and medical interns during the COVID-19 pandemic: A systematic review [J]. Nurse Educ Today. 2022;108:105183.

[21] Kailey BS, Dev D, Sado DM, et al. CardioWebinar: the evolution of digital education during the COVID-19 pandemic [J]. Heart. 2021;107(24):2004-5.

[22] Cruess R, McIlroy JH, Cruess S, et al. The Professionalism Mini-Evaluation Exercise: A Preliminary Investigation [J]. 2006;81(10):S74-S8.

[23] Cruess R, McIlroy JH, Cruess S, et al. The Professionalism Mini-Evaluation Exercise: A Preliminary Investigation [J]. Academic Medicine. 2006;81(10).

[24] Norcini JJ, Blank LL, Arnold GK, et al. The mini-CEX (clinical evaluation exercise): a preliminary investigation [J]. Ann Intern Med. 1995;123(10):795-9.

[25] Parthiban N, Boland F, Fadil Azim DH, et al. Asian medical students' attitudes towards professionalism [J]. Med Educ Online. 2021;26(1):1927466.

[26] Tsugawa Y, Ohbu S, Cruess R, et al. Introducing the Professionalism Mini-Evaluation Exercise (P-MEX) in Japan: results from a multicenter, cross-sectional study [J]. Acad Med. 2011;86(8):1026-31.

[27] Bajwa NM, Nendaz MR, Posfay-Barbe KM, et al. A Meaningful and Actionable Professionalism Assessment: Validity Evidence for the Professionalism Mini-Evaluation Exercise (P-MEX) Across 8 Years [J]. Acad Med. 2021;96(11s):S151-s7.

[28] Wu J, Wang J, Nicholas S, et al. Application of Big Data Technology for COVID-19 Prevention and Control in China: Lessons and Recommendations [J]. J Med Internet Res. 2020;22(10):e21980.

[29] Shen T, Wang C. Big Data Technology Applications and the Right to Health in China during the COVID-19 Pandemic [J]. International journal of environmental research and public health. 2021;18(14).

[30] Li X, Cong Y. A Systematic Literature Review of Ethical Challenges Related to Medical and Public Health Data Sharing in China [J]. Journal of Empirical Research on Human Research Ethics. 2021:15562646211040299.

[31] Liang W, Yao J, Chen A, et al. Early triage of critically ill COVID-19 patients using deep learning [J]. Nature communications. 2020;11(1):3543-.

[32] Meo SA, Sattar K, Ullah CH, et al. Progress and prospects of medical education research in Asian Countries [J]. Pakistan journal of medical sciences. 2019;35(6):1475-81.

[33] Schmid RD, Xiong X. Biotech in China 2021, at the beginning of the 14th five-year period ("145") [J]. Applied microbiology and biotechnology. 2021;105(10):3971-85.

[34] Wu H, Li W, Xie A, et al. Funding of health professional education: China's 20-year process and a global comparison [J]. Medical education. 2021.

[35] Silva Junior G, Askari M, Oliveira J. Digital Health and Possible Solutions to Improve the Care in the Field of Nephrology [J]. Contributions to nephrology. 2021;199:1-15.

[36] Doukas DJ, McCullough LB, Wear S. Perspective: Medical education in medical ethics and humanities as the foundation for developing medical professionalism [J]. Acad Med. 2012;87(3):334-41.

[37] Liu S, Chen Y, Li S, et al. What Are the Important Factors Influencing the Recruitment and Retention of Doctoral Students in a Public Health Setting? A Discrete Choice Experiment Survey in China [J]. International journal of environmental research and public health. 2021;18(18).

[38] Jingang A. Which future for doctors in China? [J]. Lancet. 2013;382(9896):936-7.

[39] Walkowiak D, Domaradzki J. Are rare diseases overlooked by medical education? Awareness of rare diseases among physicians in Poland: an explanatory study [J]. Orphanet journal of rare diseases. 2021;16(1):400.

[40] Domaradzki J, Walkowiak D. Knowledge and Attitudes of Future Healthcare Professionals Toward Rare Diseases [J]. Front Genet. 2021;12:639610.

Figure Legends

Figure 1. Basic knowledge for big data platform (A) before and (B) after the workshop.

* means that there is a significant difference before and after the workshop.

Figure 2. Students' attitude towards big data platform (A) before and (B) after the workshop.

* means that there is a significant difference before and after the workshop.

Figure 3. The professionalism for students (A) before and (B) after the workshop.

* means that there is a significant difference before and after the workshop.

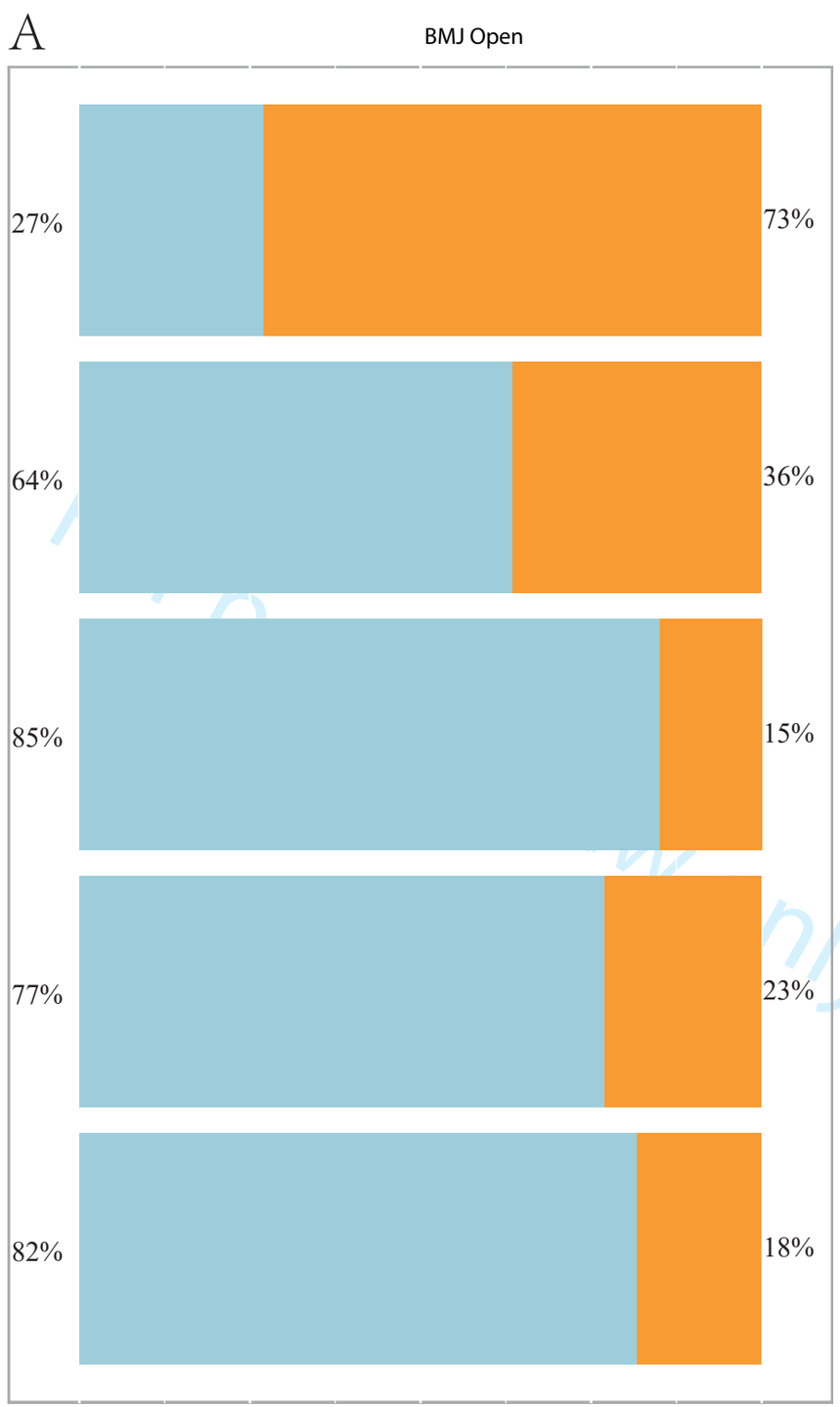
*I know the content of big data platform

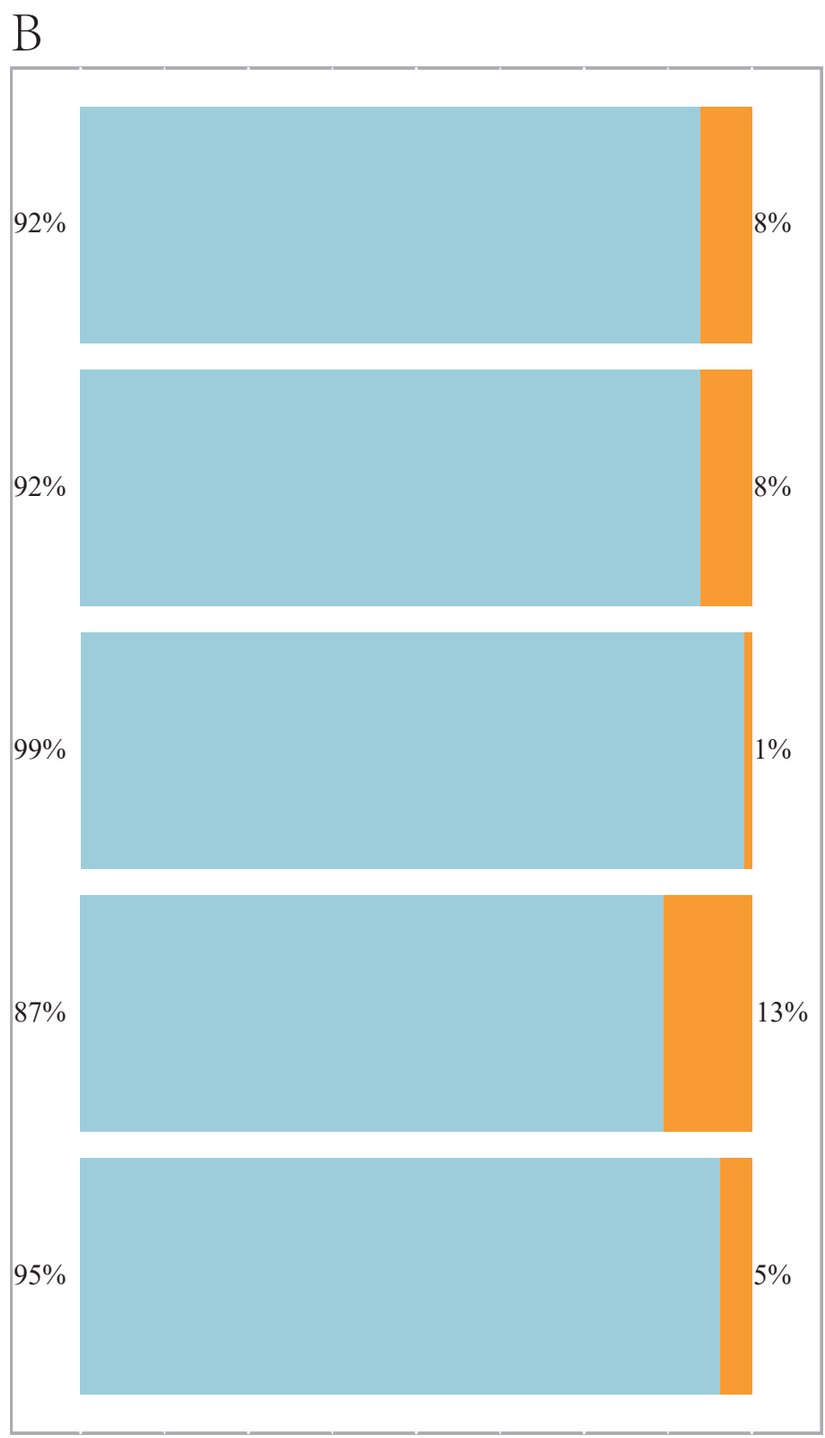
*I know the potential applications of big data technology

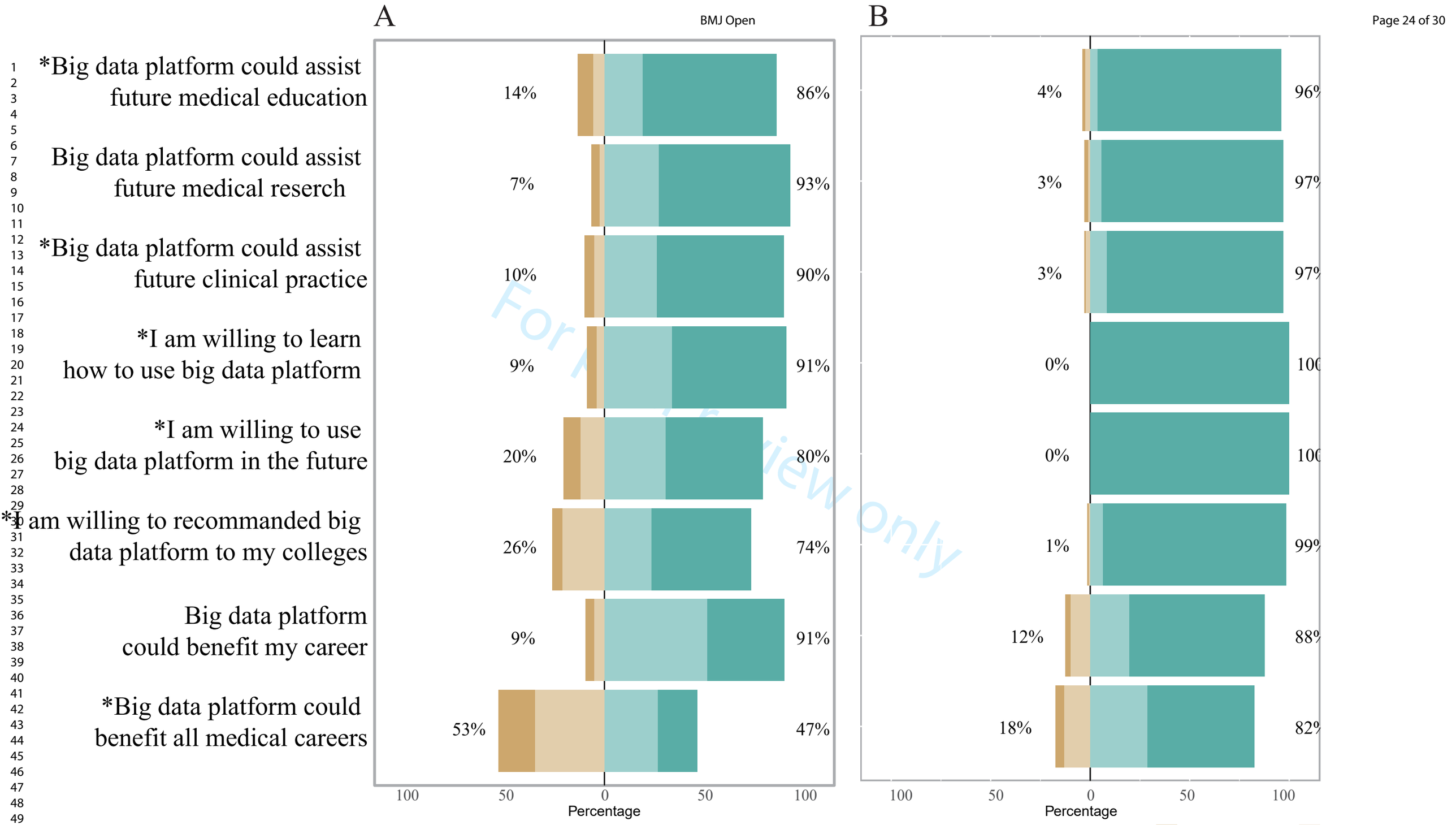
*Big data technology is practical in medical education

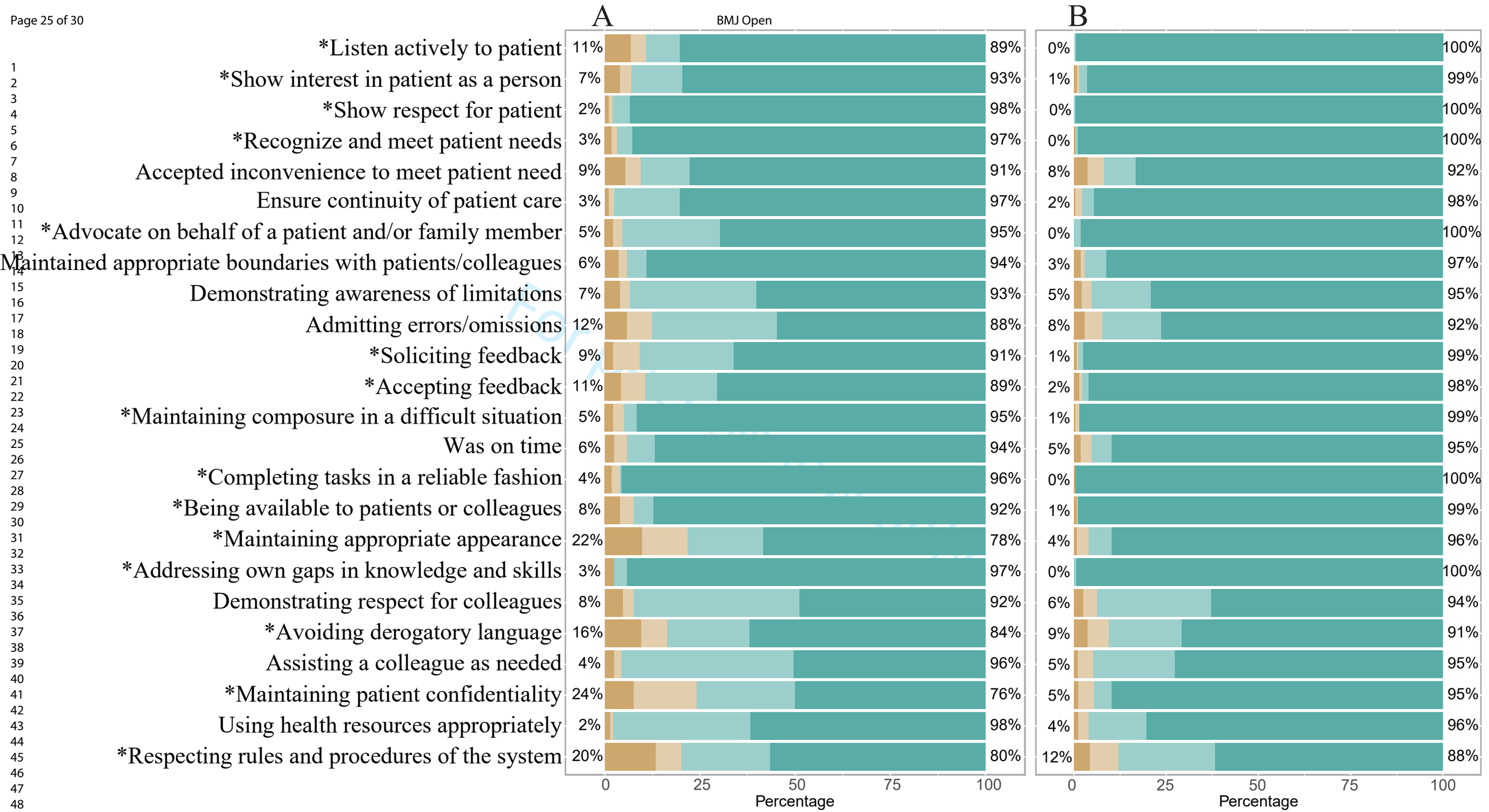
*Big data technology is practical in clinical practice

*Big data technology is practical in scientific research









Supplemental tables

Table S1, basic knowledge for big data platform before and after the workshop.

	Before Workshop		After Workshop		
Items	YES, n	NO, n	YES, n	NO, n	*P*
I know the content of big data platform	74	200	253	21	*<0.0001*
I know the potential applications of big data technology	174	100	253	21	*<0.0001*
Big data technology is practical in medical education	233	41	271	3	*<0.0001*
Big data technology is practical in clinical practice	211	63	238	36	*0.0039*
Big data technology is practical in scientific research	224	50	261	13	*<0.0001*

Table S2, students' attitude towards big data platform before and after the workshop.

	Before Workshop						After Workshop						*P*
Items	Totally disagree, n	Disagree, n	Agree, n	Totally agree, n	Mean	SD	Totally disagree, n	Disagree, n	Agree, n	Totally agree, n	Mean	SD	
Big data platform could assist future medical education	21	16	53	184	3.50	0.910	4	7	10	253	3.87	0.504	*0.0002*
Big data platform could assist future scientific research	11	7	75	181	3.55	0.735	5	3	16	250	3.86	0.499	*0.0705*
Big data platform could assist future clinical practice	13	14	72	175	3.49	0.799	2	6	23	243	3.85	0.464	*0.0017*
I am willing to learn how to use big data platform	13	11	93	157	3.43	0.783	0	0	0	274	4	0	*<0.0001*

I am willing to use big data platform in the future	23	33	84	134	3.20	0.953	0	0	0	274	4	0	*<0.0001*
I am willing to recommended big data platform to my colleagues	14	58	65	137	3.19	0.941	2	2	18	252	3.90	0.388	*<0.0001*
Big data platform could benefit my career	12	14	142	106	3.25	0.744	7	27	54	186	3.53	0.776	*0.3382*
Big data platform could benefit all medical careers	50	96	74	54	2.48	1.006	12	36	79	147	3.32	0.863	*<0.0001*

SD, Standard Deviation.

Table S3, the professionalism for students before and after the workshop.

	Before workshop				After workshop				*P*
Items	Not at all important, n	Not important, n	Important, n	Very important, n	Not at all important, n	Not important, n	Important, n	Very important, n	
Doctor-patient relationship skills									
Listen actively to patient	19	11	24	220	0	0	1	273	*<0.0001*
Show interest in patient as a person	11	8	37	218	2	2	6	264	*0.0029*
Show respect for patient	3	3	12	256	0	0	1	273	*0.0401*
Recognize and meet patient needs	5	4	11	254	1	0	2	271	*0.0255*
Accepted inconvenience to meet patient need	15	11	35	213	10	12	24	228	*0.6503*
Ensure continuity of patient care	3	4	47	220	1	5	9	259	*1*

Advocate on behalf of a patient and/or family member	6	7	70	191	0	0	5	269	*0.0008*
Maintained appropriate boundaries with patients/colleagues	10	6	14	244	5	3	16	250	*0.1440*
Reflective skills									
Demonstrating awareness of limitations	11	7	91	165	6	7	44	217	*0.4595*
Admitting errors/omissions	16	18	90	150	8	13	44	209	*0.0880*
Soliciting feedback	6	19	68	181	2	1	4	267	*<0.0001*
Accepting feedback	12	17	52	193	4	2	5	263	*0.0001*
Maintaining composure in a difficult situation	6	8	9	251	1	3	0	270	*0.0310*
Time management									
Was on time	7	9	20	238	5	8	15	246	*0.7027*
Completing tasks in a reliable fashion	5	6	1	262	1	0	0	273	*0.0086*
Being available to patients or colleagues	11	10	14	239	2	1	0	271	*0.0004*
Interprofessional relationship skills									
Maintaining appropriate appearance	27	33	54	160	2	9	17	246	*<0.0001*
Addressing own gaps in knowledge and skills	7	0	9	258	0	0	2	272	*0.0225*
Demonstrating respect for colleagues	13	8	119	134	7	10	85	172	*0.6139*
Avoiding derogatory language	26	19	59	170	10	16	54	194	*0.0220*
Assisting a colleague as needed	7	5	124	138	3	11	61	199	*0.8407*

Maintaining patient confidentiality	21	45	71	137	3	12	13	246	*<0.0001*
Using health resources appropriately	4	2	99	169	3	8	43	220	*0.3244*
Respecting rules and procedures of the system	37	18	64	155	12	21	72	169	*0.0146*

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4
Objectives	3	State specific objectives, including any prespecified hypotheses	4-5
Methods			
Study design	4	Present key elements of study design early in the paper	5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	5
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case	NA
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	6-7
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	6-7
Bias	9	Describe any efforts to address potential sources of bias	NA
Study size	10	Explain how the study size was arrived at	6
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	7-8
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	7-8
		(*b*) Describe any methods used to examine subgroups and interactions	NA
		(*c*) Explain how missing data were addressed	7-8
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	7-8
		(*e*) Describe any sensitivity analyses	NA

Continued on next page

Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	8
		(b) Give reasons for non-participation at each stage	All participants were eligible and included in this study.
		(c) Consider use of a flow diagram	NA
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	8-9
		(b) Indicate number of participants with missing data for each variable of interest	Very high response rates with no missing data, as shown in supplementary data.
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	NA
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	NA
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	NA
		Cross-sectional study—Report numbers of outcome events or summary measures	8
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	In all tables and figures.
		(*b*) Report category boundaries when continuous variables were categorized	In all tables and figures.
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	NA
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	NA
Discussion			
Key results	18	Summarise key results with reference to study objectives	13
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	15
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	16
Generalisability	21	Discuss the generalisability (external validity) of the study results	15
Other information			

Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	16

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.