Chinese experts' consensus on the application of intensive care big data

Longxiang Su[1†], Shengjun Liu[1†], Yun Long[1]*, Chaodong Chen[2], Kai Chen[3], Ming Chen[4], Yaolong Chen[5], Yisong Cheng[6], Yating Cui[7], Qi Ding[2], Renyu Ding[8], Meili Duan[9], Tao Gao[4], Xiaohua Gu[10], Hongli He[11], Jiawei He[9], Bo Hu[12], Chang Hu[12], Rui Huang[13], Xiaobo Huang[11], Huizhen Jiang[14], Jing Jiang[15], Yunping Lan[11], Jun Li[3], Linfeng Li[16], Lu Li[12], Wenxiong Li[2], Yongzai Li[17], Jin Lin[9], Xufei Luo[5], Feng Lyu[18], Zhi Mao[7], He Miao[8], Xiaopu Shang[19], Xiuling Shang[3], You Shang[20], Yuwen Shen[21], Yinghuan Shi[22], Qihang Sun[23], Weijun Sun[24], Zhiyun Tang[25], Bo Wang[6], Haijun Wang[26], Hongliang Wang[13], Li Wang[27], Luhao Wang[28], Sicong Wang[13], Zhanwen Wang[29,30,31], Zhong Wang[8], Dong Wei[22], Jianfeng Wu[29], Qin Wu[6], Xuezhong Xing[27], Jin Yang[15], Xianghong Yang[26], Jiangquan Yu[10], Wenkui Yu[4], Yuan Yu[21], Hao Yuan[28], Qian Zhai[22], Hao Zhang[26], Lina Zhang[29,30,31], Meng Zhang[15], Zhongheng Zhang[32], Chunguang Zhao[29,30,31], Ruiqiang Zheng[10], Lei Zhong[26], Feihu Zhou[7], Weiguo Zhu[33]

[1]Department of Critical Care Medicine, State Key Laboratory of Complex Severe and Rare Diseases, Peking Union Medical College Hospital, Peking Union Medical College, Chinese Academy of Medical Sciences, Beijing, China, [2]Department of Surgical Intensive Critical Unit, Beijing Chao-yang Hospital, Capital Medical University, Beijing, China, [3]Department of Critical Care Medicine, Fujian Provincial Key Laboratory of Critical Care Medicine, Shengli Clinical Medical College of Fujian Medical University, Fujian Provincial Hospital, Fujian Provincial Center for Critical Care Medicine, Fuzhou, Fujian, China, [4]Department of Critical Care Medicine, Nanjing Drum Tower Hospital, The Affiliated Hospital of Nanjing University Medical School, Nanjing, Jiangsu, China, [5]Evidence-based Medicine Center, School of Basic Medical Sciences, Lanzhou University, Lanzhou, China, [6]Department of Critical Care Medicine, West China Hospital of Sichuan University, Chengdu, China, [7]Department of Critical Care Medicine, The First Medical Center, Chinese PLA General Hospital, Beijing, China, [8]Department of Intensive Care Unit, The First Hospital of China Medical University, Shenyang, Liaoning, China, [9]Department of Critical Care Medicine, Beijing Friendship Hospital, Capital Medical University, Beijing, China, [10]Department of Critical Care Medicine, Northern Jiangsu People's Hospital; Clinical Medical College, Yangzhou University, Yangzhou, China, [11]Intensive Care Unit, Sichuan Academy of Medical Sciences & Sichuan Provincial People's Hospital, School of Medicine of University of Electronic Science and Technology, Chengdu, China, [12]Department of Critical Care Medicine, Zhongnan Hospital of Wuhan University, Wuhan, Hubei, China, [13]Department of Critical Care Medicine, The Second Affiliated Hospital of Harbin Medical University, Harbin, Heilongjiang, China, [14]Department of Information Center, Peking Union Medical College Hospital, Peking Union Medical College, Chinese Academy of Medical Sciences, Beijing, China, [15]Department of Critical Care Medicine, Chongqing General Hospital, Chongqing, China, [16]Medical Data Research Institute, Chongqing Medical University, Chongqing, China, [17]Information Network Center, QiLu Hospital, ShanDong University, Jinan, China, [18]Department of Computer Science and Engineering, Central South University, Changsha, China, [19]Department of Information Management, Beijing Jiaotong University, Beijing, China, [20]Department of Critical Care Medicine, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, [21]Intensive Care Unit of Cardiovascular Surgery Department, Qilu Hospital of Shandong University, Jinan, China, [22]National Institute of Healthcare Data Science, Nanjing University, Nanjing, China, [23]British Chinese Society of Health Informatics, Beijing, China, [24]Faculty of Automation, Guangdong University of Technology, Guangzhou, China, [25]Department of Intensive Care Unit, Zhejiang Provincial People's Hospital, Affiliated

People's Hospital, Emergency and Intensive Care Unit Center, Hangzhou Medical College, Hangzhou, Zhejiang, China, [26]Department of Intensive Care Unit, National Cancer Center/National Clinical Research Center, Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China, [27]Department of Epidemiology and Biostatistics, Institute of Basic Medical Sciences Chinese Academy of Medical Sciences; School of Basic Medicine Peking Union Medical College, Beijing, China, [28]Department of Critical Care Medicine, Sun Yat-Sen University First Affiliated Hospital, Guangzhou, China, [29]Intensive Care Unit, XiangYa Hospital, Central South University, Changsha, China, [30]National Clinical Research Center for Geriatric Disorders, Xiang Ya Hospital, Central South University, Changsha, China, [31]Hunan Provincial Clinical Research Center for Critical Care Medicine, Xiang Ya Hospital, Central South University, Changsha, China, [32]Department of Emergency Medicine, Key Laboratory of Precision Medicine in Diagnosis and Monitoring Research of Zhejiang Province, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, China, [33]Department of General Medicine, Peking Union Medical College Hospital, Peking Union Medical College, Chinese Academy of Medical Sciences, Beijing, China

The development of intensive care medicine is inseparable from the diversified monitoring data. Intensive care medicine has been closely integrated with data since its birth. Critical care research requires an integrative approach that embraces the complexity of critical illness and the computational technology and algorithms that can make it possible. Considering the need of standardization of application of big data in intensive care, Intensive Care Medicine Branch of China Health Information and Health Care Big Data Society, Standard Committee has convened expert group, secretary group and the external audit expert group to formulate Chinese Experts' Consensus on the Application of Intensive Care Big Data (2022). This consensus makes 29 recommendations on the following five parts: Concept of intensive care big data, Important scientific issues, Standards and principles of database, Methodology in solving big data problems, Clinical application and safety consideration of intensive care big data. The consensus group believes this consensus is the starting step of application big data in the field of intensive care. More explorations and big data based retrospective research should be carried out in order to enhance safety and reliability of big data based models of critical care field.

Introduction

The development of intensive care medicine is inseparable from the diversified monitoring data, which specifically presents the clinical manifestations of patients with critical symptoms. These data illustrate a certain clinical phenomena, and represents the nature of disease behind the phenomenon. Intensive care medicine has been closely integrated with data since its birth. The complexity of critical illness makes the traditional reductionist approach to medical research insufficient (1). Critical care research requires an integrative approach that embraces the complexity of critical illness and the computational technology and algorithms that can make it possible (2). Hence, the organic combination of artificial intelligence and critically ill patient data can provide significant assistance for clinical diagnosis and treatment (3). Pirracchio et al. summarize the current application of machine learning for predictive analytics and decision support in the ICU and propose online learning in the future (4). Sanchez-Pinto et al. review the definitions, types of algorithms, applications, challenges, and future of Big Data and data science

in critical care (5). There are no concenus of application of big data on the field of intensive care in China so far. Specifically, the conception, clinical research site, standard of dataset, methodology and limitation are not fully exhibited. In this experts consensus, we would like to summarize the problem above and give recommendations based on evidence.

Consensus formation

This consensus is initiated and formulated by Intensive Care Medicine Branch of China Health Information and Health Care Big Data Society, Standard Committee, and is methodically supported by the Health Data Sciences and Research Institute of Lanzhou University/Research Innovation Unit of Evidence-based Evaluation and Guidelines of Chinese Academy of Medical Sciences/Guidelines for Implementation and Knowledge Transformation Cooperation Center of the World Health Organization. This consensus has been registered on the International Practice Guide Registration Platform (Practice

guideline registration for transparency, PREPARE[1]) with the registration number being PREPARE-2022CN566.

The consensus development group consists of the consensus expert group, secretary group, working group and external audit expert group. The work flow of formation of consensus are shown in Figure 1. The enrollment criteria and obligation of these groups are shown in Supplementary material S1.

For each recommendation, external audit expert used the Likert scale (Score range: 1–6) to evaluate recommendation degree. Specifically, 6 points for total agree, 5 points for general agreement, 4 points for uncertainty, 3 points for a little disagree, 2 points for disagree and 1 point for total disagree. For each recommendation, if more than 70% external audit expert grade no less than 6 points, the consensus is reached. In this formula, 31 recommendations were put forward. Except for the two recommendations on missing value and outlier value, the remaining 29 recommendations were finalized. The degree of expert recommendation was marked with "Consensus Degree," which equal to the total number of experts no less than 6 points/total number of experts×100%.

Consensus text

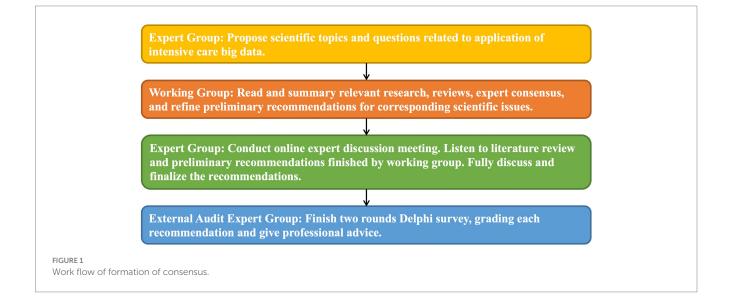### The concept, significance and necessity of intensive care data

Recommendation 1 (5–8): **The concept of intensive care big data (97% consensus)**

Intensive care big data refers to the datasets with logical connotations formulated by various indicators which are large-scale, multi-heterogeneous, variably dynamic, high-speed and real-time acquisition, low-value density and difficult to analyze traditionally in the whole process of diagnosis and treatment of patients or potential ones with critical symptoms.

---

1 http://www.guidelines-registry.org

Recommendation 2: The intensive care big data is multi-modal, massive, dynamic, continuous, and objective, and its correct acquisition can provide auxiliary evidential support for diagnosis of critical illnesses and early warning. (98% consensus)

Background and Evidence:

The monitoring methods used in the intensive care unit are abundant, and the data obtained by the combined use of multiple monitoring equipment have a multimodal characteristic (9–11). According to the needs, multi-parameter sampling can be performed at different levels and time to obtain a large amount of continuous data. Therefore, the intensive care data has the characteristics of abundance (9), dynamics, continuation, and accuracy (5, 12). Correct and effective data processing has a guiding and early-warning role in the diagnosis and treatment of critical illnesses (8). Recently, Epimed Monitor System®, a cloud-based ICU management system that includes data of more than 2.5 million hospitalization in the ICU of Brazil, has been deployed to predict the duration of ICU stays, provide guidance for risk assessment of patients becoming long-term ones in the ICU, and help to plan the use of hospital beds (13). Komorowski et al. (14) used reinforced learning techniques to guide patients with sepsis to use fluid or vasoactive medication, and external validation showed that the model made better choices for treatment than intensive care physicians. In the aspect of building predictive models by using data mining techniques, Nemati et al. (15) demonstrated that "AI sepsis experts" can be used for real-time data processing to predict new sepsis within 4–12 h. Although big data research has shown broad prospects for application, at this stage, the number of random clinical trials is small, and various technical models need to be testing prospectively in the clinic to prove their effectiveness and safety (15). In view of the characteristics of individualized and differential conditions in patients with critical symptoms (1), at this stage, the intensive care big data cannot provide maturely clinical guidance and can be used as an auxiliary support tool.

Recommendation 3: The establishment of a large database for intensive care in China should follow the principles of multiple center, multiple disease and automatic capture, and provide reliable and accurate data support for the application of big data and the development of artificial intelligence. (92% consensus)

FIGURE 1
Work flow of formation of consensus.

Background and Evidence:

The establishment of a large database for intensive care in China is in the preliminary exploration. By drawing on the experience of existing databases at home and abroad and summarizing the deficiencies of the existing databases, the database can provide a basis for clinical decision-making in China, precise medicine implementation and formulation of medical policies in China. At present, a number of large databases for intensive care have been established abroad, such as the latest version of the Intensive Care Information Database (Medical Information Mart for Intensive Care-IV, MIMIC-IV) (16), the eICU Collaborative Research Database (eICU-CRD) (17), high time resolution ICU dataset (HiRID) (18) and Amsterdam University Medical Centers Database (Amsterdam University Medical Centers Database, AmsterdamUMCdb) (19), etc., mainly based on European and American races. The volume of data is large and the types of data are abundant, but the vital signs are regularly monitored, which are not fully automatically captured, and the scoring system for critical illnesses does not have functions of automatic data collection and integration (20), there is a general lack of software embedding of preliminary data analysis online. The large database for intensive care abroad needs improving in terms of real-time data and availability. Based on the MIMIC database, the researchers conducted in-depth mining of big data to build clinical models. With the help of artificial intelligence and machine learning, artificial intelligence physicians can be used to assist clinical decision-making and provide personalized, clinical and optimal treatment for patients with critical symptoms and improve the prognosis (14).

In recent years, China has been exploring large databases for intensive care, and has successively established a database of ICU infected patients (21), a pediatric intensive care database (PIC) (22), and HeartFailure database (23), etc. The existing large databases for intensive care started late, and their development is not yet mature. They are all single-center databases, with a single type of disease or population. They are limited to the initial collection of early data, and they do not have functions of automatic data capture and data analysis, and the overall quality of data is relatively low and The utilization efficiency is not high. It has not been integrated into clinically artificial intelligence and application technology of big data (24).

Table 1 shows the brief information comparison of major foreign intensive care databases. It can be seen that the existing databases at home and abroad are mainly single-center, and various illnesses may develop into critical one and require admission to ICU for treatment, so it is significant to improve the comprehensiveness of the data. Therefore, the existing single-center or multi-center databases established for certain diseases obviously cannot meet the needs of the

vast majority of ICU patients. As we all know, the most widely used database such as MIMIC-III database records vital signs every hour, but for patients with critical symptoms who need continuous dynamic monitoring, this temporal resolution ratio is far from satisfactory. HiRID has a higher temporal resolution ratio than other published datasets, and data storage processes every 2 min (18), which is not yet possible for other databases. To sum up, the establishment of a large database for intensive care in China should follow the principles of multiple centers, multiple diseases and automatic capture to provide data support for the development and application of artificial intelligence.

Recommendation 4: Build a large database of patients with critical symptoms in China for their condition monitoring, the research and development of clinical drug and clinical trials can provide the standardized and individualized treatment for patients with critical symptoms. (97% consensus)

Background and Evidence:

Understanding the relationship between intensive care big data and critical clinic is crucial. The relationship between intensive care big data and the clinic is that: data integration can provide clinicians with manageable, interpretable, operational and treatment plan data, give certain reference to clinical treatment. Data management can provide better personalized and accurate medical guarantee through predicted and prognostic model, It can also use supervised and unsupervised learning algorithm to provide clinical researchers with handy, highly-credible and highly-utilizable database, provide scientific data support for drug development and exploration process, and finally promote the development of intensive care medicine. At present, the application of intensive care big data in clinical practice is gradually increasing, but it is mainly limited to mechanical data collection and manual data processing. The expert group believes that machine learning modeling and multi-disciplinary combination can be used to warn, track and summarize different clinical problems, so as to summarize past experience, warn current decisions and predict future progress.

The first is the application of intensive care big data in clinical decision-making. An RCT study conducted in two community hospitals in 2010 pointed out that remote data algorithms could effectively improve the medical quality of patients with critically symptoms (25). Meanwhile, a review in 2015 showed strategies for the application of big data in the use of antibiotics in patients with critically symptoms. They proposed the concept of AutoKinetics to provide decision support for clinical dosing. And through direct interaction with electronic medical records, they broadened the way to use big data and provided the right dose for each patient at the right

TABLE 1 Brief information comparison of major foreign databases of critical illnesses.

	MIMIC-III	MIMIC-IV	eICU	HiRID	Amsterdam UMCdb
Sources of included population	Single Center, Large Sample, Beth Israel Deaconess Medical Center in MIT		Multi-center, mainly small and medium-sized hospitals, organized by non-intensive specialists, with patients in 335 ICUs in the United States	Single-center, ICU patients at the University Hospital of Bern, Switzerland	Multi-center, with 20,109 ICU patients in Europe
Country/Region	USA	USA	USA	Switzerland	Europe
Time	2001–2012	2008~2019	2014~2015	2008.1~2016.6	2013~2016
Number of patients	46,520	383,220	139,367	36,098	20,109

time (26). Kindle et al. (27) and Carra et al. (8) summarized the developmental results of all remote algorithms and concluded that machine learning algorithms have important implications for sepsis detection, sepsis management, mechanical ventilation, reduction of false alarms, and prognosis in ICU. In addition, intensive care big data is also of great significance for decision making of clinical care. In 2022, the Stanford medical team developed an unsupervised process mining algorithm to evaluate the quality of care. The final result of the patient cohort had an average compliance score of 0.36. The highest was 0.64, and the lowest was 0.20. The results demonstrated the reliability of big data algorithms for data mining of electronic medical records, and the scheme could also be used to evaluate the quality of care in other diseases (28). In 2022, Jens Michael Boss et al. (29) proposed "ICU Cockpit," an integration platform of algorithmic model, which pointed out the early warning effect of severe big data on clinical decision-making. Since 2016, the platform has processed over 89 billion data points (979 patients) from 200 signals and laboratory in the analysis, and an infrastructure-based framework has been proposed for deploying and validating intensive care algorithms. It allows algorithms to seamlessly integrate into real-time data streams to generate clinical decision support and predictions in clinical practice (29). The second is the guidance of intensive care big data for clinical research. Taglang and Jackson (30) and Xu et al. (21) expounded the importance of big data to explore clinical trials systematically and, respectively. In the exploration of big data in the past 2 years, a number of studies have carried out analysis of individualized computational models constructed through big data, pointing out risk factors for high mortality in patients with critical symptoms (31–33). Finally, in terms of the relationship between clinical drug R&D and big data, we have not seen any evidence that relevant big data is used in drug R&D in the field of critical care medicine. However, due to the considerable progress of application in drug R&D and big data during recent years (34), we recommend that big data can also be combined with drug development in intensive care medicine. Therefore, the expert group recommends intensive care big data be used to detect changes in clinical practice, but more databases and algorithms and large-scale RCT experiments are needed to jointly promote the development of this field, which is also the future path of clinical practice. We point out that the multidisciplinary and interactive development of intensive care big data can build a large database of critical diseases in China, and ultimately guide the standardized treatment of patients with critical symptoms.

Clinical scientific issues concerned by intensive care big data in clinical research

Recommendation 5: It is recommended to use machine learning method to build modeling to make early warning of sepsis, acute kidney injury (AKI), and acute respiratory distress syndrome (ARDS). (94% consensus)

Background and Evidence:

Research on early warning models for sepsis, AKI, and ARDS is increasing, and most models can provide early warning with good sensitivity and specificity. The ability of different models to predict and popularize needs to be further verified. The expert group believes that machine learning method modeling can be used in the early

prediction of the risk of sepsis, AKI and ARDS in ICU patients, so as to reduce the possibility, improve early coping ability, and possibly improve prognosis.

The sepsis early warning model compared with manual screening and scoring, made early and accurate predictions, and achieved external validation. A meta-analysis of sepsis prediction models in 2020 showed that a single machine learning model can be an accurately early prediction of sepsis (AUROC 0.68–0.99) and could replace traditional scores, but heterogeneity between studies limited the evaluation of results (35). A study in 2022 (36) developed a sepsis screening tool by using a learning approach to gradient-boosted supervision that was more sensitive (84.6% vs. 80.4%) and more accurate (28.8% vs. 11.4%) than traditional scoring. A controlled study in 2021 (37) developed an algorithm that accurately predicted sepsis 12 h in advance (AUC 0.94, sensitivity 0.87, specificity 0.87). A multi-center study in 2021 (38) showed the use of a transfer-learning algorithm to enable the validity of the external validation datasets in sepsis.

Early warning models for AKI patients with critical symptoms can make early and accurate predictions, but few models have external validation, clinical interpretability, and high predictive performance in one (39). Studies have shown (40) that the early warning model of AKI has an AUC of 88%, which can predict AKI 6 h in advance. A multi-center study in 2020 showed that the AKI early warning model could predict AKI 48 h in advance, and performed well in both internal and external validation (AUC of 0.86 and 0.85, 0.86 respectively) (41). A 2020 study (42) established a continuous prediction model based on the data of electronic medical record, which could predict AKI in real time during hospitalization, and its performance was significantly better than the one-time prediction model (AUC of 0.724 vs. 0.653).

The ARDS early warning model can make early prediction of ARDS efficiently, and some models can achieve external validation, and some incorporate variables of iconography. A study in 2020 (43) using the XGBoost gradient boosting tree model could accurately predict ARDS 48 h in advance (AUROC of 79.0%). A study in 2020 (44) performed a secondary analysis of prospective study data using the text of radiology reports to build a model that performed well and achieved external validation (C-statistic, 0.78; 95% CI, 0.72–0.84). The diagnosis of ARDS is strongly dependent on iconography, which is, however, not necessarily available at the time of diagnosis or there is uncertainty in its interpretation. This information is called privileged information and uncertainty labels, but the model incorporating variables of iconography is closer to clinical practice. A study in 2021 (44) used a transfer-learning algorithm based on radiographs to build a predictive model that performed well and had external validation (AUROC of 92 and 88%). A study in 2021 (45) successfully used privileged information and a learning model with uncertainty labels to predict ARDS (AUC of 85.78 and 87.01%).

Recommendation 6: The prediction model based on machine learning can effectively predict the risk of patients at high risk of potential organ damage in the ICU. (89% consensus)

Background and Evidence:

The proposed early warning scoring system enables medical staff to better identify potential patients with critical symptoms and achieve the purpose of early identification and intervention to improve patient prognosis. However, this scoring system may fail to identify patients until significant deterioration occurs. A systematic review in 2019 (46)

found that the early warning score using statistical modeling was more accurate in identifying high-risk patients than weighted early warning (mean AUC of 0.80 vs. 0.73), with one true finding of positive case being 4.9 and 7.1 alarm events required. A similar 2021 systematic review (47) also showed that an early warning system for clinical deterioration based on machine learning could more accurately predict the risk of patient with lower survival rate in the ICU, with an area under the model ranging from 0.57 to 0.97.

Specifically, in addition to the progression of the primary disease, patients with critical symptoms may develop a variety of life-threatening comorbidities. The common ones include failure of circulatory function. In 2020, a study Hyland et al. (18) independently established an early warning system for circulatory failure, which could identify patients at risk of circulatory failure more than 2 h in advance, and successfully conducted external validation in an independent patient cohort. There was also a study by Broch Porcar et al. (48) and they considered that by using data mining, modeling, machine learning and other techniques to generate predictions, risk quantification methods could be developed to predict QTc interval prolongation. The QTc interval risk score showed good predictive performance, with good sensitivity (74% high risk, 67% intermediate risk), specificity (77% high risk, 88% intermediate risk), positive (79% high risk, 55% intermediate) and predictive value of being negative (high risk 76%, intermediate risk 88%). In addition to circulatory function and ECG function, water and electrolyte disturbances are also risk factors for patients with critical symptoms. The Spanish researchers Broch Porcar et al. (48) developed a Spanish national algorithm by reviewing the management of hyponatremia in ICU patients to improve the standardized diagnosis and treatment of hyponatremia. There was also a study (49) that the analysis group of machine learning and the analysis library of collaborative data which were based on the intensive care information system were used to know the area under the curve could be greater than 0.80 when gastrointestinal bleeding in patients was after 5 h, and it had good predictability. In addition to bleeding risk, ICU patients are also at risk of embolism. Deep vein thrombosis (DVT) is associated with high morbidity, mortality, and increased healthcare costs. Researchers (50) developed gradient boosting machine learning algorithms to predict the risk of DVT in patients 12 and 24 h before onset. The area under the curve for the diagnosis of in-hospital DVT obtained by machine learning predictors was 0.83 and 0.85, respectively.

Recommendation 7: It is recommended to use machine learning method to build modeling to conduct early screening of hospitalized patients, so as to provide help for clinicians intervene early and reduce the severity of the disease. (88% consensus)

Background and Evidence:

Compared with ordinary patients, patients with critical symptoms often undergo longer hospitalization time, more expense, and poorer prognosis. Early detection of the change of patients' condition and timely intervention are of significance for preventing the progression of the disease. Machine learning methods can facilitate early screening of diseases and timely treatment of diseases. However, for different subjects, attention should be paid to the correction of heterogeneity before the model is applied, otherwise it will easily lead to wrong clinical guidance. Experts suggest using machine learning method to building modeling for early screening of patients with critical symptoms, so as to help clinicians intervene early and reduce the severity of the disease.

A study published in 2020 evaluated several machine learning methods by using 5-fold cross validation, and applied the XGBoost algorithm to make a AI prediction model for sepsis. The validation results showed that its accuracy = 82% ± 1%; sensitivity = 65% ± 5%; specificity = 88% ± 2%; area under the receiver operating characteristic curve (AUROC) was approximately 0.89, significantly better than the SOFA score (AUROC = 0.596), which might help clinicians deploy appropriate therapeutic regimen, so early and precise responses to this AI algorithm will reduce costs, improve outcomes, and benefit healthcare systems, medical staff, and patients (51). For example, a multi-center and real-world data study in 2020 confirmed that after applying the early warning model in the clinical setting, the in-hospital mortality rate of patients with sepsis decreased by an average of 39.5%, the length of hospital stay decreased by 32.3%, and the 30-day readmission rate of sepsis-related hospitalization decreased by 22.7% (52). In addition to sepsis, machine learning methods have also been used in early screening of other critical illnesses, and a study published in 2021 used a model built with four machine learning methods (Random Forest, XGBoost, GLM-Boost, and LASSO-GLM) to predict pediatric multiple organ dysfunction (MOD). The results showed that the early prediction model of all methods achieved an AUROC of 0.91, and early prediction through risk-based patient monitoring could provide more than 22 h of lead time for the occurrence of MOD, which would play an important role in improving the prognosis of patients (53). However, there were also articles that suggest that clinicians should first calibrate the model according to the heterogeneity of patients before applying the relevant model, so as to avoid misjudgment that might affect clinical decision-making (35). However, in clinical work, first-line clinical staff should pay more attention to the existing scoring system and supervise the actual application, otherwise it will be futile to simply improve the performance of the model without improving the clinical application and response speed (54).

[Diagnosis]

Recommendation 8: It is recommended that the image data of patients with critical symptoms be included in the intensive care database to provide more comprehensive, accurate and timely diagnostic information, so as to guide clinical decision-making through relevant algorithms. (92% consensus)

Background and Evidence:

There have long been studies using AI in the screening and diagnosis of tumors and the images of infectious foci, and have confirmed its advantages in rapidly processing a large amount of image data, moved the diagnostic "gateway" forward, and avoided missed diagnosis and misdiagnosis (55, 56). The disease state and imaging manifestations of patients with critical symptoms are more complex and diverse, and the optimal timing and scenarios for using artificial intelligence for imaging diagnosis need to be more verified. The expert group believes that AI-assisted imaging diagnosis of ICU patients has good application prospects, and recommends devoting to relevant exploration to improve the efficiency and accuracy of diagnosis and provide reference for clinical decision-making.

A study of 3,078 chest radiographs from 500 ICU patients at Michigan Hospital used directional measurements and deep machine learning features to model ARDS with an accuracy of 83% and an AUC value of 0.79 (57). Cerebellar model arithmetic computer analyzed the supine chest radiograph: the AUC values for the diagnosis of pneumonia and pleural effusion were 0.737 and 0.740,

respectively, which were similar to those of imaging experts (AUC values are 0.779 and 0.698) (58). In the outbreak of COVID-19, AI-assisted imaging diagnosis has performed well. Various machine learning methods could not only quickly identify the CT images of COVID-19 (AUC values were between 0.951 and 0.980) from a large number of lung CT images, but also It could predict severe transformation in patients (AUC value was 0.848) (59). The machine learning method combining classical imaging processing and deep learning analyzed CT images of 110 patients with severe subdural hematoma, and showed that the sample recall rate and precision rate were 78.61 and 76.12%, respectively, and the specificity judged based on the severity of the hematoma volume was 92.31%, which could help physicians save decision-making time (60).

In addition to radiological imaging, AI has also been applied in other ICU bedside imaging diagnosis. One study in 2019 showed that the neural network model could detect bedside lung ultrasound B-lines with a sensitivity and specificity of 0.871 and 0.930 (61); two studies in 2021 showed that the neural network model used ultrasound images to diagnose patients with Sepsis early and the accuracy and sensitivity of developing AKI are higher than those of professional radiologists (62, 63). Electrical impedance tomography (EIT) can only roughly show the distribution of ventilation and blood flow in various regions of the lung, but it cannot be quantified as a bedside monitoring index. The neural network model trained by deep learning can calculate information such as lung volume, air flow rate, normalized airway pressure and even transpulmonary pressure from the EIT signal, and AI can also optimize the output image of EIT and even reconstruct the chest image (64).

Recommendation 9: It is recommended to divide patients with sepsis, acute kidney injury, and acute adult respiratory distress syndrome into phenotypes with different clinical outcomes and treatment responses by means of cluster analysis, and identify patients who are most likely to benefit from specific treatment strategies. (91% consensus)

Background and Evidence:

Cluster analysis can identify relatively homogeneous groups within heterogeneous populations. Some treatments are only effective in certain groups of people. Clustering techniques were used to classify patients with critical symptoms into distinct phenotypes by significant differences in comorbidities, laboratory indicators, vital signs, clinical outcomes, and treatment responsiveness, identifying groups that benefit from specific therapies. At present, the identification of phenotypes has made research progress in sepsis, AKI, and ARDS, but the accuracy and generalizability of phenotypes still need further verification. The expert group recommends that patients with critically symptoms be divided into different phenotype by cluster analysis to identify those most likely to benefit from specific treatment strategies.

Clinical and/or host response data and machine learning (e.g., latent class analysis and K-means clustering) were used to segment critically-ill patients with sepsis, AKI, ARDS, etc. into distinct phenotypes (65–68). A RCT study in 2021 identified 4 coagulation-based sepsis phenotypes by K-means clustering and used a machine learning means to determine which phenotype would benefit from rhTM (69); another RCT study by Cluster analysis identified 4 clinical phenotypes of sepsis. These phenotypes differed in demographic characteristics, laboratory abnormalities, patterns of organ dysfunction, and were not homologous to traditional patient groups

such as site of infection, pattern of organ dysfunction, or disease severity (70); a latent class analysis of an AKI cohort in 2020 identified two phenotypes of sepsis acute kidney injury with distinct clinical outcomes (71); a prospective observational cohort research through unsupervised consensus clustering and machine learning analyzed expression profiles of the whole blood RNA and identified 4 sepsis endophenotypes (Mars 1–4), of which Mars 1 was significantly associated with 28-day mortality. To facilitate clinical application, the study also extracted accurate classification biomarkers for each phenotype (72). Two different ARDS phenotypes have been identified by the LCA method using data from randomized controlled trials of ARDS. These phenotypes had different clinical outcomes. And different treatment responses to positive end-expiratory pressure strategies (73), fluid therapy (74), and simvastatin (75) have been identified.

[Treatment]

Recommendation 10: In specific clinical scenarios, such as decision making for tracheal intubation and intensive care drug decision, it is recommended to build a decision-making model that can be used for clinical treatment based on machine learning algorithms. (74% consensus)

Background and Evidence:

The condition of ICU patients is usually difficult and critical. Electronic medical record systems, monitors, ventilators and other instruments and equipment can generate massive amounts of vital information data, which far exceeds the ability of ICU doctors to continuously process and correctly interpret them, and affects the effectiveness of clinical decision-making and responsiveness. Artificial intelligence (AI) models can continuously clear, categorize, classify, calculate, and correlate a large amount of data, and make predictions for patients, thereby assisting clinical decision-making and improving the quality and efficiency of critical care.

Several studies have evaluated the clinical impact of applying artificial intelligence techniques such as machine learning to make treatment decisions. In 2018, Komorowski et al. applied reinforcement learning to the sepsis population, and AI clinicians could optimize fluid and vasoactive drug treatment and reduce the fatality rate (14). In 2019, a study established a model to predict urine output in patients with AKI. Compared with the traditional Logistic regression model, the XGBoost model could better distinguish whether patients had volume responsiveness (76).

AI technology has been tried to be applied to clinical situations such as extubation decision-making and optimization of drug treatment for patients. A 2018 retrospective study used machine learning to identify patients requiring prolonged mechanical ventilation (PMV) and those with high risk of tracheostomy (77). In 2021, Fabregat et al. compared three classification learning methods (Logistic regression, XGBoost, and support vector machines) to predict extubation outcomes, which may potentially reduce extubation failure rates (about 9%) (78). Another study in 2021 established a predictive model for accidental extubation through a machine learning algorithm, in which the random forest algorithm obtained the best AUROC of 0.787 (79).

The application of machine learning to optimize the therapeutic effects of anticoagulation, anti-infection and sedation in patients with critical symptoms is still in the exploratory stage. Chen et al. (80), Su et al. (81), Li et al. (82) compared different machine learning methods to predict the therapeutic effect of anticoagulant drugs

(citrate, heparin). The scores are overall better than the other models. A single-center retrospective study in 2022 used machine learning and cluster analysis to provide guidance on antibiotic management in patients with critical symptoms (83). Another study in 2022 based on self-attention and residual structure of convolutional neural network (CNN) had a good predictive effect on anesthesia depth monitoring (84). The examples above illustrate the potential role of AI in guiding critical decisions in patients with critical symptoms. But the vast majority of developed ICU-AI models are still in the testing or prototyping stage, and only a few have actually been evaluated in clinical practice. Van de Sande et al. found no studies suggesting the results of integrating AI models in routine clinical practice (85). Research on AI used to guide clinical decision-making is mostly calculated from retrospective and observational datasets. Therefore, in order to have AI directly guide clinical decision-making, it is necessary to conduct a comprehensive analysis of the suggested sequences or strategies derived from such AI systems with more high-quality and prospective studies to be designed.

[Prognosis and follow-up]

Recommendation 11: It is recommended to use machine learning methods to predict the prognosis of patients with critical symptoms. (85% consensus)

Background and Evidence:

There are more and more predictive models for mortality in ICU patients. Many data models are better at than disease prediction than clinical scoring systems. The sensitivity and specificity of some predictive models still rely on the assistance of clinical scoring systems. AI models in intensive care medicine are mainly generated by retrospective data, with small sample sizes and low reproducibility of conclusions, which are lack of sufficient external validation or prospective evaluation.

There are various machine learning models and algorithms, such as: support vector machines (SVM), Gradient Boosting Decision Tree (GBDT), Logistic regression (LR), adjacent algorithms (KNN, K-Nearest Neighbor), and Random Forest (RF). Studies have shown that the SVM model is a useful tool for early prediction of patients with a higher risk of death upon admission to the ICU. Compared with the early warning score of the SAPS II score, it was better at predicting 7-day mortality. However, the sensitivity and specificity of the SVM model without SAPS II significantly decreased (86). The prediction performance of the machine learning method and the traditional scoring system was further compared according to different diseases. The results were as follows: (1) Sepsis; The results in 2021 showed that GBDT is more accurate than other models (GBDT, LR, KNN, RF, and SVM) in predicting death in patients with sepsis (87). García-Gallo et al. used an assembly algorithm such as SGB to generate a sepsis model that was more accurate in predicting 1-year mortality than traditional scoring systems such as SAPS II, SOFA or OASIS (88). (2) Intracebral Hemorrhage (ICH); Nie et al. (89) indicated that RF was the best model for predicting mortality in ICH patients treated in the ICU, and all machine learning algorithms used to predict mortality in the ICU showed better results compared to the APACHE-II score. (3) Severe acute pancreatitis (SAP); Halonen et al. (90) established an artificial neural network (ANN) model for predicting the severity of acute severe pancreatitis, and the results were better than the Rason score, Glasgow-imrie, APACHE-II, and SOFA scores. The article by Ding et al. (91) also showed that the ANN

model could easily screen patients with high risk of death in the early stages of acute pancreatitis.

Finally, it is important to note that the study by Niven et al. (92) showed that a minority of critical care practices with research published in high-profile journals were evaluated for reproducibility; less than half had reproducible effects. This question highlighted the importance of accurate labeling and precise reporting methods, including data preprocessing and functionalization.

[Auxiliary decision-making system changes the clinical path]

Recommendation 12: A clinical decision support system (CDSS) can be used to improve compliance with guidelines for diagnosis and treatment of patients with critical symptoms and the implementation of clinical pathways. (86% consensus)

Background and Evidence:

Evidence-based clinical diagnosis and treatment guidelines provide standardized and homogeneous diagnosis and treatment strategies for the treatment of patients with critical symptoms. However, compliance with clinical guidelines is not high in routine ICU care, resulting in an increase in avoidable patient mortality (93, 94). A clinical decision support system (CDSS) is a computer program that helps health care workers make decisions. With the clinical application of CDSS, most studies have shown that the application of CDSS can assist ICU physicians in decision making, improve compliance with diagnosis and treatment guidelines, and improve outcomes of patient. However, there are many types of CDSSs. One CDSS is aimed at a certain disease, and the development cost is high. The CDSS based on big data has been applied to clinical decision-making, but it has not been used to change guideline compliance. Moreover, CDSS needs to be integrated with the patient electronic health record system. Due to the different electronic health record systems adopted by different regions or hospitals, the promotion and application of CDSS in different hospitals are limited. Therefore, the expert group believes that CDSS can be used to improve the compliance with the guidelines for diagnosis and treatment of patients with critical symptoms, but CDSS based on big data is still in the stage of research and development. It is recommended that qualified hospitals take the development and clinical application of CDSS based on big data into consideration to improve compliance with guidelines.

As early as in 2011, CDSS, such as a "flow sheet," can monitor various parameters of patients in real time at the bedside, screen patients with sepsis early and make a series of mandatory treatment measures according to SSC guidelines. The application of CDSS can significantly improve the compliance with SSC guideline of resuscitation bundle strategy, shorten the duration of antibiotic use (90), and reduce hospital mortality (95). In the clinical implementation of lung protective ventilation with low tidal volume, by using CDSS to guide medical staff to set the ventilator mode and support level, the compliance with lung protective ventilation improved, and the level of tidal volume increased significantly after CDSS was discontinued (96). In a study of delirium management, the duration of delirium episodes was significantly reduced, followed the adoption of the tailored ICU delirium guideline CDSS and the duration of coma was reduced, with the brain function improved (97). In another prospective observational study assessing the compliance with AKI guidelines, the CDSS for AKI was integrated into the intensive care information system in the ICU and found the proportion of patients with worse condition from stage 1 AKI, and the proportion of inappropriate use of enoxaparin dose as well as that of morbidity rate

of patients with AKI was significantly reduced (98). It can be seen that CDSS can improve guideline compliance. However, there is currently no big data-based CDSS application in clinical practice to improve guideline compliance, which needs to be confirmed by further research in the future.

Establishment, standards and principles of a large database for intensive care

Recommendation 13: It is recommended to build a intensive care medicine database and data analysis platform. (98% consensus)

Background and Evidence:

The intensive care database can provide a good data foundation and new ideas for clinical medical research, which in turn can improve the understanding of diseases. For example, in Sepsis 1.0 (99), sepsis was defined as a systemic inflammatory response syndrome (SIRS) caused by infection. Although various diagnostic indicators were more complete in Sepsis 2.0 (100), it still continued the standard of Sepsis 1.0. However, the diagnostic criteria of infection and SIRS cannot accurately describe the disease characteristics of patients, such as different primary diseases, different symptoms and mortality of patients. In 2016, Sepsis 3.0, which was mainly based on big data analysis, was born (101), which defined sepsis as life-threatening organ failure caused by the body's uncontrolled response to infection, i.e., infection and organ function diagnosis. Patterns, making the definition of Sepsis more adaptable to pathophysiology and easier to implement in clinical practice. It can be said that the intensive care medicine databases that have been constructed abroad, such as the Medical Information Mart for Intensive Care (MIMIC) and the eICU Collaborative Research Database (eICU-CRD) (17), are used in clinical practice. The role played in diagnosis and treatment has gradually become prominent. At present, the pace of establishing a intensive care big data platform has also been accelerated in China, but most of them are limited to individual databases in each hospital, and there are still some deficiencies in data exchange and influence. Therefore, we recommend building a intensive care medicine database

and data analysis for Chinese people platform to strengthen discipline construction and improve the level of treatment for patients with critical symptoms.

Recommendation 14: It is recommended to form a standard normative intensive care dataset. (97% consensus)

Background and Evidence:

Standard and normalized datasets are the basis for big data applications and facilitate data collaboration between research centers in different regions. There is a lot of information obtained by ICU equipment and instruments, and the data can be included in a reasonable and standardized manner and classified, so that they can be used more fully and conveniently. At present, there are many big data information systems for intensive care medicine at home and abroad. These information systems divide clinical data into different data elements according to specific classification standards, and then use specific data collection methods to acquire and analyze data. Referring to basic structure and data standard of the national electronic medical record (102), Beijing local standard - intensive care medicine dataset and the intensive care medicine database widely used in the field of medical research (103), the recommended standard data set should include the following data sets: (1) Basic information data of patients; (2) Diagnostic information data of patients; (3) Monitoring data of Patients; (4) Drug use data of patients; (5) Laboratory information data of patients; (6) In and out data of patients; (7) Imaging data of patients; (8) Etiology data of patients. See Table 2 for details.

It is also recommended that adjustments can be made in combination with actual conditions such as hospital disease conditions, information centers, laboratory testing items and other objective conditions. For example, based on acute respiratory distress syndrome, sepsis, acute kidney injury and other common diseases in intensive care medicine to build a special disease database, which is necessary to strengthen the sampling frequency and categories of intensive care information related to special diseases. For example, the acute respiratory distress syndrome database needs to further collection of biomarkers, etc.; The sepsis database requires further collection of vasoactive drugs, etiology collection, organ function assessment, etc.

TABLE 2 Standard datasets.

Basic information data of patients	Time information on patient admission and discharge, demographic information, source of admission, ICU category, time of death, etc.
Diagnostic information data of patients	All disease diagnosis information during the patient's stay in the ICU; the main diagnosis needs to be distinguished from the secondary diagnosis
Monitoring data of Patients	Routine vital signs, ventilator parameter information, blood purification parameter information, aortic balloon counter pulsation parameter information, the mental state, the score information, etc.
Treatment data of patients	The route of administration, use time and drug dose of all drugs during the patient's stay in the ICU; the name, time and related information of the operation; the name, time and related information of the treatment operation, etc.
Laboratory information data of patients	Laboratory examination information during the patient's stay in the ICU, such as sampling time, specimen type, test items, reference range of normal values, etc.
In and out data of patients	Data of all fluids entering and expelling from the body during the patient's stay in the ICU, including fluid type, entry and exit route, time, etc.
Imaging data of patients	Text reports related to radiographic imaging during patient stay in the ICU
Etiology data of patients	The etiological data collected during the patient's stay in the ICU, including sampling time, specimen type, etiological name, etiological drug susceptibility, etc.

Recommendation 15: It is recommended to select automatic collection for objective data first. For data that cannot be automatically collected for the time being, targeted collection should be carried out in combination with research needs, data sources and data types. (92% consensus)

Background and Evidence:

The data collection process should follow the principles of comprehensiveness, multi-dimensionality, efficiency and timeliness. In view of the many data sources and rich data structures in the ICU, it is recommended to use automated data collection technology to realize the data collection process so as to avoid human errors affecting the use of subsequent data.

Data in the ICU can be broadly classified into "phenotypic data" and "physiologic data." Phenotypic data include demographics, age, sex, laboratory values, and physician and nursing records. Phenotypic data collection can be queried and extracted from electronic medical records (EMRs). Relevant content can be obtained through Python or API, and the required attribute content can be extracted from it. Physiological data include vital signs (blood pressure, heart rate, respiratory rate, core temperature) and other parameters (intracranial pressure, EEG) generated by bedside monitoring equipment. If the data interface of the device can be obtained through various software manufacturers, data collection and aggregation can be realized through the interface docking method. If some devices cannot obtain the data interface, collecting all the data generated by the target device can be tried to acquire the underlying data exchange of the system, the network package between the client and the database, which can convert the data into with restructure and output to new database, based on underlying IO request and network analysis technologies.

Alarms in the ICU, such as ECG leads, blood pressure cuff detachment from patients, completion of infusion pump or air bubbles in tubing, high airway pressure, air leak, or apnea in mechanical ventilation ventilators, etc., which can be classified into the type of physiological data. This part of the data can be collected by collecting logs from log sources of various devices. Continuous waveform data is more complicated to acquire due to its continuous nature and high sampling rate. In recent years, many studies have used time series databases and unstructured databases such as InfluxDB, MongoDB, etc. to explore the writing, storage, and query processes of various continuous-time signals, which can solve the storage-transmission-exchange-exploitation problem (104). For image data, since most of the images are currently stored in the PACS system, it is necessary to clarify whether to collect from the equipment (CT machine, ultrasound machine, etc.) or through the PACS docking port (105).

Recommendation 16: It is recommended to optimize standard system for intensive care big data, standardize multi-center source data, and constrain standard codes, measurement units, field standards, as well as naming dictionaries to ensure the homogeneity and standardization of the use of the large database for intensive care. (95% consensus)

Background and Evidence:

"Information integration, standards first" (106), the construction of large databases for intensive care must be implemented in accordance with the corresponding norms and standards, the standard codes, measurement units, field standards, and naming dictionaries, and it is constrained by standard norms to ensure the subsequent modeling and application process. The consistency of data processing ensures the standardized production of data from the source, and lays

the foundation for the construction, data integration, data exchange and data sharing of large databases for intensive care. Intensive care big data are multi-modal data with high privacy and diverse sources, and have the characteristics of multiple data dimensions, good timeliness, high value density and high data quality. The "phenotypic data" and "physiological data" in the ICU can be classified into structured discrete data, time series data, and unstructured text data, image data, and audio-video data (107). The main contents are as follows: (1) Discrete data: basic information and routine data of patients' physical sign, including a series of discrete data such as gender, age, blood type, height, weight, etc., which are mainly characterized data. These data volumes are small and stable. (2) Time series data: mainly physiological data, including time series data of various vital sign parameters such as blood oxygen, heart rate, and ECG. These data are closely related to the real-time symptoms of patients, with high real-time performance, strong continuity, and large datasets. (3) Image data: mainly physiological data, including a large amount of image data such as ultrasound and radiation. These image data are large in volume and are important reference data for diagnosis and operation. (4) Text data: a large amount of text data about patient medical records and diagnostic results, mainly for representation data, including electronic medical records, surgical records, inspection reports, etc. Among all data types of critical diseases, time series data, image data and text data have high information value density and play an important role in clinical diagnosis, treatment and decision making.

Due to the uneven level of informatization in each center and a wide range of coverage, the above-mentioned data formats for intensive care are complicated and difficult to integrate. After negotiation, multiple centers have formulated unified data fields, contents and formats for the big database for intensive care, and established a standard system. For example, for the standardization of image data, the level of imaging departments in different hospitals varies, and multiple centers need to negotiate the image quality standards for uploading compressed original images. For different types of data, in order to ensure the standardization of large databases for intensive care, data governance rules for different types of data can be formulated, and the system will automatically clean the data when it enters the database, supplemented by manual review if necessary to ensure data quality. For the quality assessment of inbound data, it can be measured from normative (the extent to which the data conforms to data standards, data models, business rules, metadata or authoritative reference data), integrity (the extent to which data elements are assigned values according to data rules), accuracy (the degree to which the data accurately represents the true value of the real entity, "real object" that it describes), consistency (the degree to which the data does not contradict the data used in other specific contexts), timeliness (the degree to which the data is correct over time), and accessibility (the degree to which data can be accessed), which are six aspects to manage and evaluate (108).

Recommendation 17: It is recommended to establish a data security system to ensure the security of data storage, processing, sharing and use. (98% consensus)

Background and Evidence:

The information security system in China mainly includes five technical tasks: risk assessment and grade protection, monitoring system, cryptography and network trust system, emergency response system, and disaster preparedness. The security level of

information system is divided into five levels, and the levels from one to five are gradually increased. Centering on the "Network Security Law," "Data Security Law" and "Personal Information Protection Law," China carries out the construction of data classification system. In terms of data security, the security of the data itself (using modern cryptographic algorithms to actively protect data) and security of data protection (active protection of data using modern information storage methods) must both be paid attention to. New security issues need to be addressed in an environment of big data, including balancing privacy and utility, analyzing and governing encrypted data, and verifying authenticated and anonymous users. With the continuous expansion of the application scope of intensive care big data, the content is becoming richer and more valuable with a large amount of sensitive personal information. A security system and a safety management responsibility system for intensive care big data should be established to ensure the security in data storage, opening, and processing (109, 110).

When storing data, system security reinforcement as well as software and hardware architecture design in a distributed environment (such as Apache Hadoop) should be done well. Strict fine-grained access control and risk registration management strategies should be set for static data, and privacy-related data storage should realize classified isolation data encryption (such as AES, RSA, SHA-256 and other encryption methods) and other security technical means, dynamic data classification and identification of important sensitive data should be through encryption and dynamic audit capabilities, using TLS (transport layer security technology) to communicate between cluster nodes and maintain confidentiality during transmission, and enabling unified management across platforms (endpoints, mobile devices, networks, and storage systems) (106).

During data processing, the software architecture and network configuration should be designed according to the database volume and access method, especially for multi-center, and the appropriate hardware architecture should be designed according to the software architecture. And policy configuration such as network security should be done to ensure data security. After the data is authorized to be processed by other parties, the most important question is whether there is misuse and malicious restoration of sensitive data during the processing, whether it complies with laws and regulations, and whether it complies with the privacy clauses agreed by both parties (106). In multi-party computation, data leakage is avoided through system policy design such as data desensitization (111) and federated learning (112).

When sharing data, measures such as data desensitization, rights management, and log auditing should be taken to ensure data security. Data cannot be unconditionally open to the public or third parties. Consideration should be given to the fact that sensitive information can be easily restored after a single information is desensitized through multi-source collisions which may lead to security risks, therefore, only point-to-point sharing, or multilateral transactions based on certain special constraints, such as sharing health records, patient medication information, medical images and other information about intensive care big data. Whether the data sharing is justified or not should be comprehensively weighed on the occasions of the data and the subject's right to know.

Ways and methods to solve big data problems in intensive care medicine

[Type of data]

Recommendation 18: It is recommended to use processing methods of digital signals such as filters to preprocess time series data, deep learning to process image data, use Natural Language Processing (NLP) technology to process unstructured text data. (93% consensus)

Background and Evidence:

From the perspective of machine model building, intensive care data can be roughly divided into four categories: numerical time series data, numerical non-series data, text data, and image data. Among them, numerical data can be divided into two categories according to the collection density: (1) time series data, or "streaming data," including electrocardiogram, arterial and intracranial pressure, hemodynamic monitoring, ventilator data, brain waves and other data with relatively high collection frequency; (2) non-sequential data, or "sparse data," including blood gas analysis, laboratory test results, medical history and other data with relatively low collection frequency. Different types of data can be combined to improve the accuracy of AI prediction models (113), provide decision support under complex and uncertain diagnostic conditions (114), and better adapt to the clinical real-time data environment.

For time series data, before further pattern recognition or other processing through different algorithms, processing methods of digital signals such as filters are usually used for preprocessing. The main purpose is to use various mathematical methods to strip components of different frequencies in the signal for targeted treatment. For example, in electrocardiogram (ECG) data processing, a five-minute moving average is often used for low-pass and high-pass filtering (29, 115, 116), and when building an EEG signal model, Narula et al. also used a band-pass filter to remove baseline drift and high-frequency interference (117).

For non-series data, the processing skills are mainly reflected in solving the problems of data (parameter) outliers and missing values, screening and dimensionality reduction according to different algorithm models. After the corresponding preprocessing of the data, whether it is a simple algorithm such as linear regression and logistic regression, or a sophisticated algorithm such as lifting algorithm and reinforcement learning (14, 118), it can achieve good results in the corresponding scene. So no special recommendation is made.

For image data, such as CT, pathological slices, ultrasound images, etc., most of them are processed by deep learning (such as convolutional neural network CNN, etc.) and other tasks (119–122). In particular, Walsh et al. believed that deep learning methods can directly extract important features from images, which could help to generate novel biomarkers and more accurate image analysis tools (123).

For unstructured text data, such as narrative text in EMR, as well as radiology, pathology reports, etc., the content can be mined and processed through natural language processing technology to obtain pathological information, social environment information, etc., which can be combined with the existing expert knowledge base (such as the unified medical language system, etc.) as a supplement to improve the accuracy of related prediction models, and show a speed and accuracy that exceeds manual processing (124–126). In particular, natural language processing for Chinese, ICTCLAS system, THULAC toolkit,

etc. are all good auxiliary tools, but the Sinicization of knowledge bases such as UMLS (or other Chinese medical knowledge bases) needs to be demonstrated in the literature.

[Data preprocessing]

Recommendation 19: It is recommended to use resampling methods to deal with unbalanced datasets. (78% consensus)

Background and Evidence:

In intensive care medical datasets, unbalanced data is very common. Unbalanced data refers to the uneven distribution of the number of samples among each category in the classification task, and there will be a particularly large gap, which will greatly affect the final performance of the prediction model. For example, a small number of death samples in intensive care medicine datasets carry important information about mortality prediction, but are ignored because the model is insensitive to data imbalances. In response to the phenomenon of data imbalance, the expert group recommends using resampling methods to process imbalanced data, which are mainly divided into three types: undersampling, oversampling and synthetic oversampling. Undersampling is the random sampling of fewer samples from most classes so that the data tends to be balanced. Edited Nearest Neighbors (ENN) is the most typical undersampling method. Oversampling is to generate more labeled samples according to the sample rules with fewer sample labels so that the data tend to be balanced. Synthetic Minority Over-sampling Technique (SMOTE) is an oversampling technique that generates synthetic samples for the minority class. In order to reduce the fitting problem caused by oversampling and undersampling, a method combining oversampling and undersampling is extended to deal with data imbalance on this basis, such as SMOTEENN, SMOTETomek, etc. In the study of using machine learning to predict atrial fibrillation, Tiwari et al. used a variety of sampling methods to deal with the imbalance problem that the data in the control group was much more than that in the experimental group, and compared the data under different sampling methods, and finally chose the random oversampling method according to the classifier effect (127). Papp et al. used SMOTE sampling to synthesize samples from the minority class for the class-imbalance problem and analyzed the synthesized new data results through cross validation and confusion matrix (128).

Recommendation 20: It is recommended to convert original categorical variables and numerical variables into variables that can be directly processed by machine learning algorithms through one-hot encoding, sequential encoding, etc. (83% consensus)

Background and Evidence:

The function of variable category transformation is to convert the original category of intensive care medical data containing the above information into a form suitable for data mining and easy for model understanding. The transformation of variable categories makes the original data more tidy and consistent through operations such as encoding. It is recommended to use methods such as one-hot encoding and sequential encoding. One-hot encoding is a common numerical processing method for unordered categorical variables, with "1" to indicate that it belongs to this category, and "0" to indicate that it does not belong to this category. One-hot encoding will add new variables to the original variables. The number of new variables being added is the number of types. Ordinal coding is a common numerical processing method for ordinal categorical variables. This coding makes numerical one according to the different degrees represented by

the ordinal variables, such as scores about a patient's health status from 0 to 5.

Recommendation 21: It is recommended to use dimensionality reduction methods such as principal component analysis to perform variable screening of high-dimensional features in intensive care datasets. (90% consensus)

Background and Evidence:

In most research problems of intensive care big data, the datasets used usually have high-dimensional feature variables, which can easily lead to overfitting problems and increase training costs. Therefore, it is necessary to extract important features through variable screening to achieve the purpose of data dimensionality reduction. Experts recommend principal component analysis, variance selection, univariate feature selection, regularization models, feature ranking based on machine learning models, and recursive feature elimination methods.

Principal Component Analysis (PCA) is a popular general feature dimensionality reduction method, which can be used to reduce the dimensionality of various types of data such as numerical values, texts, and images. Essentially, multiple variables are synthesized into a few independent components, and each component can reflect the information of the original variable, which can improve the learning speed and reduce the training cost. Variance selection is a simple feature selection method that filters features by removing features with low variance. Univariate feature selection usually uses statistical test methods such as chi-square test and F test, or measures such as Pearson correlation coefficient and distance correlation coefficient to determine the relationship between variables. The regularization model is mainly divided into L1 regularization and L2 regularization. By adding additional constraints or penalty terms to the loss function of the existing model, it can prevent overfitting and improve the generalization ability of the model. L2 regularization is more stable than L1 regularization and is more favorable for the understanding of features. Regularization models are often used in feature selection of medical data. In the study on early triage of COVID-19 patients with critical symptoms, Liang et al. selected 10 statistically significant variables as predictors by the Lasso method (129). Many machine learning methods can achieve feature scoring, such as feature ranking by measuring feature importance. Therefore, it is recommended to use the selected machine learning model to complete feature selection, including SVM, random forest, decision tree, XGBoost, LGBM and other models. By adjusting the calculation parameters of feature importance, the feature ranking of different methods can be obtained. This method is convenient, effective and easy to understand the relationship between the model and features, but it is needed to verify the model fitting effect by means of cross-validation. In addition, recursive feature elimination methods can be considered to screen the features of intensive care medical data.

[Model Construction]

Recommendation 22: It is recommended to select supervised learning, unsupervised learning and reinforcement learning models for critical disease prediction and identification according to different scenarios and different data types. (97% consensus)

Background and Evidence:

The intensive care unit monitoring system collects a large number of the patients' respiratory, hemodynamic, neurological and clinical data, and its electronic medical record system also records the patient's

clinical treatment and medication information in detail. The data types include types of text, digital and image. Through the processing and analysis capabilities of big data by machine learning algorithms, key features of the data can be mined to assist in diagnosis and decision making. Machine learning algorithms can be classified into supervised learning, unsupervised learning, and reinforcement learning, depending on whether the dataset has labels. Among them, supervised learning can learn and summarize models, including decision trees, support vector machines, random forests, naive Bayesian models, artificial neural networks and other models; Unsupervised learning models can discover hidden patterns without manual annotation or data grouping, which can find potential similarities and differences in the data. Common algorithms include k-means, principal component analysis, hierarchical clustering, etc.; Reinforcement learning can learn the best behavior or mode that should be taken from experience. The model type should be selected according to the data type and medical task. Among them, for numerical data and clinical prediction problems, supervised learning models can be used; For text data, natural language processing models and unsupervised learning models can be used; For image data, Convolutional Neural Networks (CNN), Recurrent Neural Network (RNN) can be used for medical image recognition and segmentation; For clinical auxiliary decision-making tasks, reinforcement learning models can be used. According to the literature survey, the usage scenarios of three different learning methods include: (1) Supervised learning: prognosis prediction, phenotype classification, analgesic and sedation strategy selection, mortality risk prediction, disease severity prediction, prediction for length of stay in the ICU, etc. (2) Unsupervised learning: disease pattern mining and representation based on electronic health records (EHR). (3) Reinforcement learning: decision making of treatment plan, recommendation of fluid volume, robot-assisted surgery, etc.

Specifically, examples of the usages and indications for the three types of learning are as follows:

1 Supervised learning: Prognosis prediction and dose recommendation for heparin patients (82); monitoring and adjustment of Local citric acid anticoagulation (80); prediction of in-hospital mortality risk in patients with critical symptoms (124, 130) prediction of mortality risk in patients with candidemia (125), prediction of the severity of lung ultrasound in ICU patients (126), etc.

2 Unsupervised learning: Phenotype classification and sedation strategy selection in mechanically ventilated patients (131); temperature pattern recognition in patients with critical symptoms (67), blood pressure pattern recognition (132); subtype of diseases extracted from electronic health record data (133, 134).

3 Reinforcement learning: Dynamically provide optimal treatment plan and select intravenous fluids and vasopressor doses for patients in the ICU (135).

Recommendation 23: It is recommended to use a causal inference model to explore and discover causal relationships in the intensive care field. (89% consensus)

Background and Evidence:

The model system of causal inference is built on the basis of causal-heuristic learning and reasoning. It conducts in-depth mining of relevant data to extract causal structure, and conducts causal-heuristic estimation. It studies the influence of intervention variables on prognosis and obtains the key index of prognosis evaluation. The directions involved include causal discovery, causal structure learning, causal inference, causal deep learning, etc. In response to the need for poor ICU prognosis or poor survival rate, as well as the need to accurately determine the influencing factors of prognosis, it is advisable to use the frameworks of DoWhy, CDT and CausalML and establish a causal-heuristic learning inference and decision-making evaluation system based on the database of specialized diseases and multi-center of intensive care.

First, implement big data-driven causal structure identification, mine causal relationships, and conduct feature analysis, effect analysis, and interpretability analysis. Richens et al. (136) proposed a counterfactual diagnostic strategy for expected failure and expected adequacy, breaking the traditional diagnosis method of diseases based on symptoms and narrowing the scope of possible conditions by using counterfactual questions. Wei et al. (137) described the causal relationship between some variables in the recommended system from the perspective of causal inference and solved the influence of popularity bias on the model from the perspective of counterfactual inference. Goudet et al. (138) used deep learning methods to propose a causal generative neural network (CausalGNN), which exploited conditional independence and distribution asymmetry to discover bivariate and multivariate causal structures, and learned functional causal models from observational data to figure out a causal roadmap between clinicopathological features.

In addition, the causal effect was further estimated on the basis of the causal relationship, and machine learning methods such as generalized random forest (GRF) (139) were used to calculate CATE and HTE to predict the difference in prognosis under different ICU intervention methods and research the degree of impact on prognosis by intervention variables. Tan et al. (140) used an approach like adversarial training to give an interpretable means for recommended systems. The advantages of these methods are that the data can be used to reason about the source characteristics of heterogeneity to estimate a series of estimators, which also apply to high-dimensional data and missing data and have good interpretability. Through techniques based on causal discovery and estimation, learning the most discriminative characterization, discovering diagnostic basis and key characteristic indexes, judging prognosis accurately, and providing effective interventions for clinical treatment can be realized.

[**Verification of the model**]

Recommendation 24: It is recommended to add external validation to internal validation of the model. (94% consensus)

Background and Evidence:

Model validation is the process of evaluating the predictive performance of a model after it has been constructed. The importance of model validation is reflected in measuring the prediction accuracy of the prediction model, feeding back the model building process, and adjusting the model building ideas if necessary. The model verification idea is relatively mature at present, and there is a relatively consistent method consensus. In practice, model verification is mainly divided into internal verification and external verification. The expert group believes that the following methods can be used to evaluate the model validation process.

Internal verification: In general, verification based on their own data (internal verification) is required. That is, a part of the data (like

80% of the total) are randomly selected as the training set for building the prediction model, and the rest of the data are used as the test set to evaluate the performance of the model. In order to verify that the model has good performance on newly generated clinical data, "spatio-temporal division" can be added to the random division, which is the data of the latest period specially divided as an independent validation set (141). In order to improve the estimation robustness of the evaluation indicators, K-fold cross-validation can be used (18). Divide the data set into K parts (such as 10 parts), use K-1 data to build a prediction model, use the remaining data for verification, repeat K times, and take the average of the K times of model prediction evaluation indicators as the accuracy index of the final model. The implementation of internal verification is relatively simple, but since the training set and test set are both derived from the same data, the model extrapolation ability (i.e., "generalization" ability) is relatively weak.

External verification: Different regions and hospitals may have differences in data distribution due to differences in population, disease characteristics, and diagnosis and treatment habits. In order to verify that the model has good extrapolation, it is recommended to perform external validation on multi-center data from different regions and different hospitals.

Recommendation 25: It is recommended to use indicators such as sensitivity, specificity, F1 score, and AUC to evaluate the performance of classification models, and indexes such as R^2, MSE, RMSE, and MAE to evaluate the performance of regression models. (91% consensus)

Background and Evidence:

During model validation, a series of evaluation metrics should be used to measure model performance (i.e., predictive effect). For classification model and regression model, different indicators are used for evaluation.

Performance evaluation indicators of classification model: For classification models (models whose predicted values are categorical variables), sensitivity (also known as recall), specificity, F1 score, precision, AUC (Area Under Curve) and other metrics to evaluate the performance are generally used (3). Among them, the F1 score is the harmonic value of sensitivity and positive accuracy rate, and the larger the value, the better the model performance is. AUC is the area under the ROC curve drawn by "1-specificity" and "sensitivity." The larger the value, the better the model performance is. When the sample categories are not balanced, it is recommended to use the area under the PR curve, AUPRC, to evaluate model performance.

Performance evaluation indicators of regression model: For regression models (models whose predicted values are continuous variables), R^2 (R squared, coefficient of determination, coefficient of determination), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and other indicators are generally used to evaluate the performance. The closer the determinant coefficient, R^2, is to 1, the better the model performance is. The closer MSE, RMSE and MAE are to 0, the better the model performance is.

[**Model interpretability**]

Recommendation 26: It is recommended to explore the interpretability of the model to facilitate the clinical transformation of complex machine learning models. The recommended model interpretation methods include Feature Importance, LIME, and Shapley. (91% consensus)

Background and Evidence:

AI models based on big data training in intensive care medicine are often complex, and their complexity is mainly reflected in the large number of parameters and the complex functional relationship between various parameters. Such complex models are often not conducive to clinicians to analyze the pathophysiological mechanisms, and it is difficult to determine the causal relationship between variables, which seriously hinders the clinical transformation of AI research results. The interpretability of the model is considered as an effective way to solve the above problems. Understanding characteristics, classification, and prediction of indicators, and then understanding why a machine learning model makes such a decision, and what features play the most important role in the decision allow us to judge whether the model is in line with common sense. For example, an AI doctor trained by a reinforcement learning model is used to treat septic shock (14). The AI prompts the need to increase norepinephrine while appropriately limiting fluid replacement. Understanding the mechanism behind such an algorithm is critical for the reliability of the model. If the algorithm tells you that you need to increase the dose of norepinephrine for the patient because their main contradiction is peripheral vasodilation, rather than fluid deficiency, it can greatly enhance the confidence of the physician in the use of this model, because the diagnosis and treatment made by AI decision-making is consistent with clinical pathophysiological changes.

In addition, several other methods are also used for model interpretability exploration (142). Feature importance can be used. Its main working principle is to change the arrangement of the data in a certain column of the data table and keep the rest of the features unchanged to see how much it affects the prediction accuracy.

Locally Interpretable Agnostic Modeling (LIME) is an algorithm (143) that provides a novel technique to interpret the results of any predictive model in an interpretable and trustworthy way. It works by training an interpretable model locally around the predictions it wants to explain. In layman's terms, select a sample and a point near the sample, and then train a simple model to fit. Although the simple model cannot be effective on the complete data set, it is at least effective near this point. The characteristics of this simple model are human-analyzable, and the trained weights can also represent feature importance.

The Shapley value was proposed by Loyd Shapley, a professor at the University of California, Los Angeles, USA, to solve the problem of contribution and profit distribution in cooperative games. In cooperation of N persons, the contribution of individual member is different, and the distribution of income should also be different. The ideal distribution method is: contribution = income; Is there a quantifiable method for the distribution of contribution and income? The Shapley method is one such method, where the Shapley value of a feature is the average marginal contribution of that feature across all feature sequences.

Clinical application of intensive care big data

Recommendation 27: It is recommended to transform and promote early warning tools that meet critical needs. (91% consensus)

Background and Evidence:

The construction of early warning tools can make early predictions for the risk of various adverse events in the ICU, thus helping clinicians to take timely measures to prevent problems before they occur, reduce the incidence of adverse events in patients effectively, and improve early response capabilities. At present, although early warning models have been constructed and verified for the occurrence and prognosis of a variety of critical diseases at home and abroad, there is still insufficient research to truly conduct large-scale clinical trials to evaluate their application value. A Big-data clinical trial (BCT) of an early warning tool was implemented in terms of injury and disease deterioration. However, there are still differences in the predictive performance of different early warning tools in different application scenarios, and further promotion and verification are needed. So far, no mature disease-targeted early warning tools have been launched at home and abroad. The expert group believes that it is possible to use AI technology to provide early warning for various adverse events in the ICU. At the same time, it is necessary to carry out BCT research to further verify the clinical practical value of early warning tools, so as to achieve early detection, diagnosis and treatment of diseases.

For the early warning of sepsis, Shimabukuro et al. (144) conducted a BCT study in 2017 and found that patients who used the early warning tool for sepsis shortened the length of hospital stay significantly (10.3 days vs. 13.0 days, $p = 0.042$), and the in-hospital mortality rate was reduced significantly (8.96% vs. 21.3%, $p = 0.018$). However, a single-center BCT study conducted by Semler et al. (145) found that the application of a sepsis electronic warning system neither improved the completion of the 6-h bundle of sepsis ($p = 0.159$), nor improved clinical outcomes (including ICU fatality rate, days in ICU, days of vasoactive drug use).

For the early warning of acute kidney injury (AKI), a large multi-center BCT study in the United States in 2021 (146) found that the AKI early warning system did not improve disease progression in patients ($p = 0.67$). However, BCT evidence from the United Kingdom (147) and China (148) found that although the AKI early warning system could not improve the mortality rate of patients, it could significantly improve the early identification rate of AKI (RR: 1.12, 95% CI: 1.03–1.22, $p < 0.01$) and AKI diagnosis rate (7.9% vs. 2.7%, $p = 0.001$). Another BCT study from the United States found that the AKI electronic automatic alert system did not improve the composite outcome (maximum creatinine change, the need for dialysis or death) within 7 days of patients ($p = 0.88$) (149).

For the early warning of disease deterioration, the Escobar et al. (150) conducted a multi-center BCT study in 2020 that included a total of 43,949 people (15,487 people in the intervention group and 28,462 people in the control group). And it found that early warning tool for disease progression can significantly reduce patient mortality rate (adjusted RR: 0.84, 95% CI: 0.78–0.90, $p < 0.001$).

Recommendation 28: It is recommended to use the information system for intensive care as a carrier to access real-time data and output recommendations for decision making. (91% consensus)

Background and Evidence:

The condition of patients with critical symptoms is complex and fast-changing, and ICU equipment and instruments have a large amount of information, so the data dimension and the update frequency is high. The application carrier should be effectively integrated with the hospital information system, which can obtain high-dimensional information in real time, and be equipped with a

prediction model. Based on Hadoop distributed processing technology, Xia et al. (33) designed a big data analysis system for intensive care medicine, and conducted a performance test through the "Study on the Effect of Xuebijing on AKI-related Sepsis" (33). The information system of intensive care big data can integrate ICU high-dimensional information, obtain analysis data in real time, and use it as a carrier for results of intensive care big data such as prediction models and scores (151). Boss et al. (29) developed an online real-time ICU decision support platform that could be used to collect multimodal waveform data and AI-based computational disease modeling, calling it "ICU Cockpit." In the cohort of 979 patients admitted to this 12-bed neurocritical care unit since 2016, the total number of data points processed and stored by the "ICU Cockpit" platform has been approximately 88.9 billion (29). Based on the intensive care information system, Zhang Suzhen et al. used the XGBoost model to integrate relevant parameters and performed machine learning to predict the risk of AKI in patients with septic shock. The sensitivity of the prediction results was 73.3%, the specificity was 71.7%, and the accuracy was 72.5%. Compared with the traditional score, it was significantly improved (152).

When there is no information system of intensive care medicine, the intensive care big data can also be equipped with a online prediction tool of web page, APP, applet, or bedside form and other carriers. Flechet et al. developed a prediction model for acute kidney injury, AKI predictor, and conducted a multicenter prospective cohort study to verify the prediction effect of clinicians and AKI Predictor. The performance of the two at ICU admission was: AUROC was 0.80 [0.69–0.92] and 0.75 [0.62–0.88] ($n = 120$, $p = 0.25$), the net benefit ranges were 0–26% and 0–74%. The machine learning-based AKI predictor achieved similar discriminative performance to physicians in predicting AKI-2 and AKI-3, with a higher overall net benefit because physicians overestimated the risk of AKI. This showed that AKI Predictor has added value to the doctor's prediction. The study also came with an online version of the predictive model[2] (153).

With the development of Internet of Things technology, 5G technology, database technology, etc., the carrier to realize the application of intensive care big data in the future should focus more on the "dynamic holographic prediction system" that obtains ICU information in a comprehensive and real-time way, analyzes the data and makes real-time predictions.

Recommendation 29: It is suggested that the current practice of intensive care diagnosis and treatment should still be led by clinicians with the use of big data technology to coordinate to improve medical efficiency and ensure medical quality and safety. (98% consensus)

Background and Evidence:

In recent years, the development of applications of intensive care big data has made rapid progress, and a large number of articles have been published, including prediction of diseases, early warning of risks, and real-time guidance of clinical medication. In the foreseeable future, big data applications can assist ICU clinical diagnosis and treatment activities. However, at the same time, applications of big data still have problems such as lack of clinical integration, lack of high-quality verification, poor interpretability, few application

---

2   https://www.akipredictor.com/en/

TABLE 3 Summary of recommendations on the application of intensive care big data.

Part	No.	Recommendation contents	Consensus degree
1. The concept, significance and necessity of intensive care data	1	The concept of intensive care big data: Intensive care big data refers to the datasets with logical connotations formulated by various indicators which are large-scale, multi-heterogeneous, variably dynamic, high-speed and real-time acquisition, low-value density and difficult to analyze traditionally in the whole process of diagnosis and treatment of patients or potential ones with critical symptoms	97
	2	The intensive care big data is multi-modal, massive, dynamic, continuous, and objective, and its correct acquisition can provide auxiliary evidential support for diagnosis of critical illnesses and early warning	98
	3	The establishment of a large database for intensive care in China should follow the principles of multiple center, multiple disease and automatic capture, and provide reliable and accurate data support for the application of big data and the development of artificial intelligence	92
	4	Building a large database of patients with critical symptoms in China for their condition monitoring, the research and development of clinical drug and clinical trials can provide the standardized and individualized treatment for patients with critical symptoms	97
2. Clinical scientific issues concerned by intensive care big data in clinical research	5	It is recommended to use machine learning method to build modeling to make early warning of sepsis, acute kidney injury (AKI), and acute respiratory distress syndrome (ARDS)	94
	6	The prediction model based on machine learning can effectively predict the risk of patients at high risk of potential organ damage in the ICU	89
	7	It is recommended to use machine learning method to build modeling to conduct early screening of hospitalized patients, so as to provide help for clinicians intervene early and reduce the severity of the disease	88
	8	It is recommended that the image data of patients with critical symptoms be included in the intensive care database to provide more comprehensive, accurate and timely diagnostic information, so as to guide clinical decision-making through relevant algorithms	92
	9	It is recommended to divide patients with sepsis, acute kidney injury, and acute adult respiratory distress syndrome into phenotypes with different clinical outcomes and treatment responses by means of cluster analysis, and identify patients who are most likely to benefit from specific treatment strategies	91
	10	In specific clinical scenarios, such as decision making for tracheal intubation and intensive care drug decision, it is recommended to build a decision-making model that can be used for clinical treatment based on machine learning algorithms	74
	11	It is recommended to use machine learning methods to predict the prognosis of patients with critical symptoms	85
	12	A clinical decision support system (CDSS) can be used to improve compliance with guidelines for diagnosis and treatment of patients with critical symptoms and the implementation of clinical pathways	86
3. Establishment, standards and principles of a large database for intensive care	13	It is recommended to build a intensive care medicine database and data analysis platform	98
	14	It is recommended to form a standard normative intensive care dataset	97
	15	It is recommended to select automatic collection for objective data first. For data that cannot be automatically collected for the time being, targeted collection should be carried out in combination with research needs, data sources and data types	92
	16	It is recommended to establish a standard system for intensive care big data, standardize multi-center source data, and constrain standard codes, measurement units, field standards, as well as naming dictionaries to ensure the homogeneity and standardization of the use of the large database for intensive care	95
	17	It is recommended to establish a data security system to ensure the security of data storage, processing, sharing and use	98

(Continued)

TABLE 3 (Continued)

Part	No.	Recommendation contents	Consensus degree
4. Ways and methods to solve big data problems in intensive care medicine	18	It is recommended to use processing methods of digital signals such as filters to preprocess time series data, deep learning to process image data, use Natural Language Processing (NLP) technology to process unstructured text data	93
	19	It is recommended to use resampling methods to deal with unbalanced datasets	78
	20	It is recommended to convert original categorical variables and numerical variables into variables that can be directly processed by machine learning algorithms through one-hot encoding, sequential encoding, etc.	83
	21	It is recommended to use dimensionality reduction methods such as principal component analysis to perform variable screening of high-dimensional features in intensive care datasets	90
	22	It is recommended to select supervised learning, unsupervised learning and reinforcement learning models for critical disease prediction and identification according to different scenarios and different data types	97
	23	It is recommended to use a causal inference model to explore and discover causal relationships in the intensive care field	89
	24	It is recommended to add external validation to internal validation of the model	94
	25	It is recommended to use indicators such as sensitivity, specificity, F1 score, and AUC to evaluate the performance of classification models, and indexes such as R^2, MSE, RMSE, and MAE to evaluate the performance of regression models	91
	26	It is recommended to explore the interpretability of the model to facilitate the clinical transformation of complex machine learning models. The recommended model interpretation methods include Feature Importance, LIME, and Shapley	91
5. Clinical application of intensive care big data	27	It is recommended to transform and promote early warning tools that meet critical needs	91
	28	It is recommended to use the information system for intensive care as a carrier to access real-time data and output recommendations for decision making	91
	29	It is suggested that the current practice of intensive care diagnosis and treatment should still be led by clinicians with the use of big data technology to coordinate to improve medical efficiency and ensure medical quality and safety	98

scenarios, and ethics. Therefore, this consensus believes that because of the current developmental level of big data applications, it is advisable to be guided by existing evidence and clinical experience to assist the diagnosis and treatment, and improve the quality and efficiency of medical care with the help of big data technology.

Big data models produce seemingly accurate results through complex computations, but often fail to provide end users with the logic behind them (154). AI is weak in determining causality, at least its interpretability does not meet current clinical needs. Models developed based on intensive care big data are often more accurate in predictions when validating data from the same population, but the results may be unreliable when tested in external populations (155). In clinical practice, the diagnosis and treatment process is often highly subjective, especially for patients with critically complex symptoms, and their plans for diagnosis and treatment also have large individual heterogeneity, resulting in low reliability of the ICU model (156). In summary, most of the current research is still in the exploited phase and lacks effective external validation (157). Therefore, unnecessary interventions or changes in treatment strategies that are not supported by scientific evidence may lead to medical safety issues such as overmedication or treatment failure.

When these algorithms are developed into intelligent assistance systems deployed as alerting tools, they must be concise and accurate enough to prevent alert fatigue and thus avoid delays in clinical

decision-making (158, 159). Considering the scientific preciseness, the maturity and stability of AI-driven models are less convincing for clinical practice to a certain extent, and indiscriminate development and use of data models may lead to overdiagnosis and waste of resources. In addition, the clinical application of intensive care big data also faces ethical issues. At present, the hidden dangers of big data applications in terms of patient privacy and safety responsibility cannot be ignored. First of all, the establishment of the database will inevitably involve data of patient privacy, and protecting patient privacy has become a problem that must be solved in the development of intensive care big data. It is not appropriate to develop a medical database at full speed without guaranteeing privacy and security. Secondly, in terms of application security, in the process of big data-assisted clinical diagnosis and treatment practice, if a medical safety accident occurs, computer algorithms cannot be responsible for clinical decision-making with the current developmental level of ethics and AI. In order to avoid mistakes and abuses in the big data system for diagnosis and treatment, the clinician must act as the person in charge of clinical decision-making to "be responsible for" big data applications.

Discussion

With the increase of computing power and data scale, the emergence of large models has enabled AI systems to handle more complex and

massive tasks, improving the model's performance and generalization ability, which also brings new opportunities for critical big data applications. As a "double-edged sword," the application of big data science in intensive care has pros and cons. This consensus reach a consensus on five parts: conception, important scientific issues, standards and principles of database, methodology in solving big data problems, clinical application and safety consideration of intensive care big data. All recommendations has been summarized in Table 3. Actually, this is the starting step of application big data in the field of intensive care. In order to ensure data security and ensure the professionalism of the model, the medical industry needs a medical vertical domain big language model based on professional mapping knowledge domain and high-quality data. More explorations and big data based retrospective research should be carried out in order to enhance safety and reliability of big data based models of critical care field.

Author contributions

LS and SL: write and translate the Chinese version of this consensus and overall planning of the discussion and finalize the draft. YLo: organize and supervise the writing of this article. CC, KC, MC, YCheng, YCui, QD, TG, XG, HH, JH, CH, RH, HJ, JJ, YLa, JuL, LuL, JiL, XL, ZM, HM, YuS, QS, WS, ZT, HaW, LuW, SW, ZhaW, ZhoW, DW, QW, JYu, YY, HY, HZ, MZ, CZ, RZ, and LeZ: literature review and write the origin version of the Chinese version of this consensus. YChen, RD, MD, BH, XH, LiL, WL, YLi, FL, XiaS, XiuS, YoS, YiS, BW, HoW, LiW, JW, XX, JYa, XY, WY, QZ, LiZ, ZZ, FZ, and WZ: supervise and provide consulting of manuscript. All authors contributed to the article and approved the submitted version.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2023.1174429/full#supplementary-material

References

1. Celi LA, Mark RG, Stone DJ, Montgomery RA. Big data in the intensive care unit. Closing the data loop. *Am J Respir Crit Care Med*. (2013) 187:1157–60. doi: 10.1164/rccm.201212-2311ED

2. Buchman TG, Billiar TR, Elster E, Kirk AD, Rimawi RH, Vodovotz Y, et al. Precision medicine for critical illness and injury. *Crit Care Med*. (2016) 44:1635–8. doi: 10.1097/CCM.0000000000002028

3. Jung Y, Hu J. A K-fold averaging cross-validation procedure. *J Nonparamet Stat*. (2015) 27:167–79. doi: 10.1080/10485252.2015.1010532

4. Pirracchio R, Cohen MJ, Malenica I, Cohen J, Chambaz A, Cannesson M, et al. Big data and targeted machine learning in action to assist medical decision in the ICU. *Anaesth Crit Care Pain Med*. (2019) 38:377–84. doi: 10.1016/j.accpm.2018.09.008

5. Sanchez-Pinto LN, Luo Y, Churpek MM. Big data and data science in critical care. *Chest*. (2018) 154:1239–48. doi: 10.1016/j.chest.2018.04.037

6. Rumsfeld JS, Joynt KE, Maddox TM. Big data analytics to improve cardiovascular care: promise and challenges. *Nat Rev Cardiol*. (2016) 13:350–9. doi: 10.1038/nrcardio.2016.42

7. Yang S, Stansbury LG, Rock P, Scalea T, Hu PF. Linking big data and prediction strategies: tools, pitfalls, and lessons learned. *Crit Care Med*. (2019) 47:840–8. doi: 10.1097/CCM.0000000000003739

8. Carra G, Salluh JIF, da Silva Ramos FJ, Meyfroidt G. Data-driven ICU management: using big data and algorithms to improve outcomes. *J Crit Care*. (2020) 60:300–4. doi: 10.1016/j.jcrc.2020.09.002

9. Le Roux P, Menon DK, Citerio G, Vespa P, Bader MK, Brophy GM. Consensus summary statement of the international multidisciplinary consensus conference on multimodality monitoring in Neurocritical care: a statement for healthcare professionals from the Neurocritical care society and the European Society of Intensive Care Medicine. *Neurocrit Care*. (2014) 21:S1–S26. doi: 10.1007/s12028-014-0041-5

10. Citerio G, Park S, Schmidt JM, Moberg R, Suarez JI, Le Roux PD. Data collection and interpretation. *Neurocrit Care*. (2015) 22:360–8. doi: 10.1007/s12028-015-0139-4

11. Schmidt JM, De Georgia M. Multimodality monitoring: informatics, integration data display and analysis. *Neurocrit Care*. (2014) 21:S229–38. doi: 10.1007/s12028-014-0037-1

12. Docherty A. B and N I lone, exploiting big data for critical care research. *Curr Opin Crit Care*. (2015) 21:467–72. doi: 10.1097/MCC.0000000000000228

13. Zampieri FG, Soares M, Borges LP, Salluh JIF, Ranzani OT. The Epimed monitor ICU database®: a cloud-based national registry for adult intensive care unit patients in Brazil. *Revista Brasileira de Terapia Intensiva*. (2017) 29:418–26. doi: 10.5935/0103-507X.20170062

14. Komorowski M, Celi LA, Badawi O, Gordon AC, Faisal AA. The artificial intelligence clinician learns optimal treatment strategies for sepsis in intensive care. *Nat Med*. (2018) 24:1716–20. doi: 10.1038/s41591-018-0213-5

15. Nemati S, Holder A, Razmi F, Stanley MD, Clifford GD, Buchman TG. An interpretable machine learning model for accurate prediction of Sepsis in the ICU. *Crit Care Med*. (2018) 46:547–53. doi: 10.1097/CCM.0000000000002936

16. Johnson ABL, Pollard T, Horng S, Celi LA, Mark R. MIMIC-IV (version 2.0). *PhysioNet*. (2022)

17. Pollard TJ, Johnson AEW, Raffa JD, Celi LA, Mark RG, Badawi O. The eICU collaborative research database, a freely available multi-center database for critical care research. *Sci Data*. (2018) 5:180178. doi: 10.1038/sdata.2018.178

18. Hyland SL, Faltys M, Huser M, Lyu X, Gumbsch T, Esteban C. Early prediction of circulatory failure in the intensive care unit using machine learning. *Nat Med*. (2020) 26:364–73. doi: 10.1038/s41591-020-0789-4

19. Thoral PJ, Peppink JM, Driessen RH, Sijbrands EJG, Kompanje EJO, Kaplan L, et al. Sharing ICU patient data responsibly under the Society of Critical Care Medicine/European Society of Intensive Care Medicine Joint Data Science Collaboration: the Amsterdam university medical centers database (AmsterdamUMCdb) example. *Crit Care Med*. (2021) 49:e563–77. doi: 10.1097/ccm.0000000000004916

20. Qi S, Mao Z, Hu X, Liu C, Kang H, Zhou F. Introduction of critical care database based on specialized information systems: a model of critical care medicine database in large level III grade a hospital. *Zhonghua Wei Zhong Bing Ji Jiu Yi Xue*. (2020) 32:743–9. doi: 10.3760/cma.j.cn121430-20200520-00393

21. Xu P, Chen L, Zhu Y, Yu S, Chen R, Huang W, et al. Critical care database comprising patients with infection. *Front Public Health*. (2022) 10:852410. doi: 10.3389/fpubh.2022.852410

22. Zeng X, Yu G, Lu Y, Tan L, Wu X, Shi S, et al. PIC, a paediatric-specific intensive care database. *Sci Data*. (2020) 7:14. doi: 10.1038/s41597-020-0355-4

23. Zhang Z, Cao L, Chen R, Zhao Y, Lv L, Xu Z, et al. Electronic healthcare records and external outcome data for hospitalized patients with heart failure. *Sci Data*. (2021) 8:46. doi: 10.1038/s41597-021-00835-9

24. Xu J, Zhou Y, Xia X. Construction and application of big data platform for intensive care medicine. *Chin J Emerg Med*. 31:129–132. doi: 10.3760/cma.j.issn.1671-0282.2022.01.028

25. Morrison JL, Cai Q, Davis N, Yan Y, Berbaum ML, Ries M, et al. Clinical and economic outcomes of the electronic intensive care unit: results from two community hospitals. *Crit Care Med*. (2010) 38:2–8. doi: 10.1097/CCM.0b013e3181b78fa8

26. Elbers PW, Girbes A, Malbrain ML, Bosman R. Right dose, right now: using big data to optimize antibiotic dosing in the critically ill. *Anaesth Intens Ther*. (2015) 47:457–63. doi: 10.5603/AIT.a2015.0061

27. Kindle RD, Badawi O, Celi LA, Sturland S. Intensive care unit telemedicine in the era of big data, artificial intelligence, and computer clinical decision support systems. *Crit Care Clin*. (2019) 35:483–95. doi: 10.1016/j.ccc.2019.02.005

28. Noshad M, Rose CC, Chen JH. Signal from the noise: a mixed graphical and quantitative process mining approach to evaluate care pathways applied to emergency stroke care. *J Biomed Inform*. (2022) 127:104004. doi: 10.1016/j.jbi.2022.104004

29. Boss JM, Narula G, Straessle C, Willms J, Azzati J, Brodbeck D, et al. ICU cockpit: a platform for collecting multimodal waveform data, AI-based computational disease modeling and real-time decision support in the intensive care unit. *J Am Med Inform Assoc*. (2022) 29:1286–91. doi: 10.1093/jamia/ocac064

30. Taglang G, Jackson DB. Use of "big data" in drug discovery and clinical trials. *Gynecol Oncol*. (2016) 141:17–23. doi: 10.1016/j.ygyno.2016.02.022

31. Zhu Y, Yin H, Zhang R, Ye X, Wei J. The effect of dobutamine vs milrinone in sepsis: a big data, real-world study. *Int J Clin Pract*. (2021) 75:e14689. doi: 10.1111/ijcp.14689

32. Huang X, Shan S, Khan YA, Salem S, Mohamed A, Attia EA. Risk assessment of ICU patients through deep learning technique: a big data approach. *J Glob Health*. (2022) 12:04044. doi: 10.7189/jogh.12.04044

33. Xia Y, Wang X, Wu W, Shi H. Rehabilitation of Sepsis patients with acute kidney injury based on intelligent medical big data. *J Healthc Eng*. (2022) 2022:8414135–10. doi: 10.1155/2022/8414135

34. Vergetis V, Skaltsas D, Gorgoulis VG, Tsirigos A. Assessing drug development risk using big data and machine learning. *Cancer Res*. (2021) 81:816–9. doi: 10.1158/0008-5472. Can-20-0866

35. Fleuren LM, Klausch TLT, Zwager CL, Schoonmade LJ, Guo T, Roggeveen LF, et al. Machine learning for the prediction of sepsis: a systematic review and meta-analysis of diagnostic test accuracy. *Intensive Care Med*. (2020) 46:383–400. doi: 10.1007/s00134-019-05872-y

36. Delahanty RJ, Alvarez J, Flynn LM, Sherwin RL, Jones SS. Development and evaluation of a machine learning model for the early identification of patients at risk for Sepsis. *Ann Emerg Med*. (2019) 73:334–44. doi: 10.1016/j.annemergmed.2018.11.036

37. Goh KH, Wang L, Yeow AYK, Poh H, Li K, Yeow JJL, et al. Artificial intelligence in sepsis early prediction and diagnosis using unstructured data in healthcare. *Nat Commun*. (2021) 12:711. doi: 10.1038/s41467-021-20910-4

38. Wardi G, Carlile M, Holder A, Shashikumar S, Hayden SR, Nemati S. Predicting progression to septic shock in the emergency department using an externally generalizable machine-learning algorithm. *Ann Emerg Med*. (2021) 77:395–406. doi: 10.1016/j.annemergmed.2020.11.007

39. Ozrazgat-Baslanti T, Loftus TJ, Ren Y, Ruppert MM, Bihorac A. Advances in artificial intelligence and deep learning systems in ICU-related acute kidney injury. *Curr Opin Crit Care*. (2021) 27:560–72. doi: 10.1097/MCC.0000000000000887

40. Xiao Z, Huang Q, Yang Y, Liu M, Chen Q, Huang J, et al. Emerging early diagnostic methods for acute kidney injury. *Theranostics*. (2022) 12:2963–86. doi: 10.7150/thno.71064

41. Churpek MM, Carey KA, Edelson DP, Singh T, Astor BC, Gilbert ER, et al. Internal and external validation of a machine learning risk score for acute kidney injury. *JAMA Netw Open*. (2020) 3:e2012892. doi: 10.1001/jamanetworkopen.2020.12892

42. Kate RJ, Pearce N, Mazumdar D, Nilakantan V. A continual prediction model for inpatient acute kidney injury. *Comput Biol Med*. (2020) 116:103580. doi: 10.1016/j.compbiomed.2019.103580

43. Le S, Pellegrini E, Green-Saxena A, Summers C, Hoffman J, Calvert J, et al. Supervised machine learning for the early prediction of acute respiratory distress syndrome (ARDS). *J Crit Care*. (2020) 60:96–102. doi: 10.1016/j.jcrc.2020.07.019

44. Mayampurath A, Churpek MM, Su X, Shah S, Munroe E, Patel B, et al. External validation of an acute respiratory distress syndrome prediction model using radiology reports. *Crit Care Med*. (2020) 48:e791–8. doi: 10.1097/CCM.0000000000004468

45. Jabbour S, Fouhey D, Kazerooni E, Wiens J, Sjoding MW. Combining chest X-rays and electronic health record (EHR) data using machine learning to diagnose acute respiratory failure. *J Am Med Informat Assoc*. (2022) 29:1060–8. doi: 10.1093/jamia/ocac030

46. Linnen DT, Escobar GJ, Hu X, Scruth E, Liu V, Stephens C. Statistical modeling and aggregate-weighted scoring Systems in Prediction of mortality and ICU transfer: a systematic review. *J Hosp Med*. (2019) 14:161–9. doi: 10.12788/jhm.3151

47. Muralitharan S, Nelson W, Di S, McGillion M, Devereaux PJ, Barr NG, et al. Machine learning-based early warning Systems for Clinical Deterioration: systematic scoping review. *J Med Internet Res*. (2021) 23:e25187. doi: 10.2196/25187

48. Broch Porcar MJ, Cubillo BR, Domínguez-Roldán JM, Rocha LÁ, Sanz MÁB, Montes MC, et al. Practical document on the management of hyponatremia in critically ill patients. *Med Intensiva (Engl Ed)*. (2019) 43:302–16. doi: 10.1016/j.medin.2018.12.002

49. Levi R, Carli F, Arevalo AR, Altinel Y, Stein DJ, Naldini MM, et al. Artificial intelligence-based prediction of transfusion in the intensive care unit in patients with gastrointestinal bleeding. *BMJ health care*. *Inform*. (2021) 28:e100245. doi: 10.1136/bmjhci-2020-100245

50. Ryan L, Mataraso S, Siefkas A, Pellegrini E, Barnes G, Green-Saxena A, et al. A machine learning approach to predict deep venous thrombosis among hospitalized patients. *Clin Appl Thromb Hemost*. (2021) 27:1076029621991185. doi: 10.1177/1076029621991185

51. Yuan KC, Tsai LW, Lee KH, Cheng YW, Hsu SC, Lo YS, et al. The development an artificial intelligence algorithm for early sepsis diagnosis in the intensive care unit. *Int J Med Inform*. (2020) 141:104176. doi: 10.1016/j.ijmedinf.2020.104176

52. Burdick H, Pino E, Gabel-Comeau D, McCoy A, Gu C, Roberts J, et al. Effect of a sepsis prediction algorithm on patient mortality, length of stay and readmission: a prospective multicentre clinical outcomes evaluation of real-world patient data from US hospitals. *BMJ Health Care Inform*. (2020) 27:e100109. doi: 10.1136/bmjhci-2019-100109

53. Bose SN, Greenstein JL, Fackler JC, Sarma SV, Winslow RL, Bembea MM. Early prediction of multiple organ dysfunction in the pediatric intensive care unit. *Front Pediatr*. (2021) 9:711104. doi: 10.3389/fped.2021.711104

54. Bedoya AD, Clement ME, Phelan M, Steorts RC, O'Brien C, Goldstein BA. Minimal impact of implemented early warning score and best practice alert for patient deterioration. *Crit Care Med*. (2019) 47:49–55. doi: 10.1097/CCM.0000000000003439

55. Ardila D, Kiraly AP, Bharadwaj S, Choi B, Reicher JJ, Peng L, et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nat Med*. (2019) 25:954–61. doi: 10.1038/s41591-019-0447-x

56. Qin ZZ, Sander MS, Rai B, Titahong CN, Sudrungrot S, Laah SN, et al. Using artificial intelligence to read chest radiographs for tuberculosis detection: a multi-site evaluation of the diagnostic accuracy of three deep learning systems. *Sci Rep*. (2019) 9:15000. doi: 10.1038/s41598-019-51503-3

57. Reamaroon N, Sjoding MW, Gryak J, Athey BD, Najarian K, Derksen H. Automated detection of acute respiratory distress syndrome from chest X-rays using directionality measure and deep learning features. *Comput Biol Med*. (2021) 134:104463. doi: 10.1016/j.compbiomed.2021.104463

58. Rueckel J, Kunz WG, Hoppe BF, Patzig M, Notohamiprodjo M, Meinel FG, et al. Artificial intelligence algorithm detecting lung infection in supine chest radiographs of critically ill patients with a diagnostic accuracy similar to board-certified radiologists. *Crit Care Med*. (2020) 48:e574–83. doi: 10.1097/CCM.0000000000004397

59. Zhang K, Liu X, Shen J, Li Z, Sang Y, Wu X, et al. Clinically applicable AI system for accurate diagnosis, quantitative measurements, and prognosis of COVID-19 pneumonia using computed tomography. *Cells*. (2020) 181:1423–1433.e11. doi: 10.1016/j.cell.2020.04.045

60. Farzaneh N, Williamson CA, Jiang C, Srinivasan A, Bapuraj JR, Gryak J, et al. Automated segmentation and severity analysis of subdural hematoma for patients with traumatic brain injuries. *Diagnostics (Basel, Switzerland)*. (2020) 10:773. doi: 10.3390/diagnostics10100773

61. van Sloun RJG, Demi L. Localizing B-lines in lung ultrasonography by weakly supervised deep learning, *in-vivo* results. *IEEE J Biomed Health Inform*. (2020) 24:957–64. doi: 10.1109/JBHI.2019.2936151

62. Lv Y, Huang Z. Account of deep learning-based ultrasonic image feature in the diagnosis of severe Sepsis complicated with acute kidney injury. *Comput Math Methods Med*. (2022) 2022:8158634–9. doi: 10.1155/2022/8158634

63. Ying F, Chen S, Pan G, He Z. Artificial intelligence pulse coupled neural network algorithm in the diagnosis and treatment of severe Sepsis complicated with acute kidney injury under ultrasound image. *J Healthc Eng*. (2021) 2021:1–8. doi: 10.1155/2021/6761364

64. Strodthoff N, Strodthoff C, Becher T, Weiler N, Frerichs I. Inferring respiratory and circulatory parameters from electrical impedance tomography with deep recurrent models. *IEEE J Biomed Health Inform*. (2021) 25:3105–11. doi: 10.1109/JBHI.2021.3059016

65. Bos LD, Schouten LR, van Vught LA, Wiewel MA, Ong DSY, Cremer O, et al. Identification and validation of distinct biological phenotypes in patients with acute respiratory distress syndrome by cluster analysis. *Thorax*. (2017) 72:876–83. doi: 10.1136/thoraxjnl-2016-209719

66. Chaudhary K, Vaid A, Duffy Á, Paranjpe I, Jaladanki S, Paranjpe M, et al. Utilization of deep learning for subphenotype identification in Sepsis-associated acute kidney injury. *Clin J Am Soc Nephrol*. (2020) 15:1557–65. doi: 10.2215/CJN.09330819

67. Bhavani SV, Carey KA, Gilbert ER, Afshar M, Verhoef PA, Churpek MM. Identifying novel Sepsis subphenotypes using temperature trajectories. *Am J Respir Crit Care Med*. (2019) 200:327–35. doi: 10.1164/rccm.201806-1197OC

68. Sinha P, Churpek MM, Calfee CS. Machine learning classifier models can identify acute respiratory distress syndrome phenotypes using readily available clinical data. *Am J Respir Crit Care Med*. (2020) 202:996–1004. doi: 10.1164/rccm.202002-0347OC

69. Kudo D, Goto T, Uchimido R, Hayakawa M, Yamakawa K, Abe T, et al. Coagulation phenotypes in sepsis and effects of recombinant human thrombomodulin: an analysis of three multicentre observational studies. *Crit Care*. (2021) 25:114. doi: 10.1186/s13054-021-03541-5

70. Seymour CW, Kennedy JN, Wang S, Chang CCH, Elliott CF, Xu Z, et al. Derivation, validation, and potential treatment implications of novel clinical phenotypes for Sepsis. *JAMA*. (2019) 321:2003–17. doi: 10.1001/jama.2019.5791

71. Wiersema R, Jukarainen S, Vaara ST, Poukkanen M, Lakkisto P, Wong H, et al. Two subphenotypes of septic acute kidney injury are associated with different 90-day mortality and renal recovery. *Crit Care*. (2020) 24:150. doi: 10.1186/s13054-020-02866-x

72. Scicluna BP, van Vught LA, Zwinderman AH, Wiewel MA, Davenport EE, Burnham KL, et al. Classification of patients with sepsis according to blood genomic endotype: a prospective cohort study. *Lancet Respir Med*. (2017) 5:816–26. doi: 10.1016/s2213-2600(17)30294-1

73. Calfee CS, Delucchi K, Parsons PE, Thompson BT, Ware LB, Matthay MA, et al. Subphenotypes in acute respiratory distress syndrome: latent class analysis of data from two randomised controlled trials. *Lancet Respir Med*. (2014) 2:611–20. doi: 10.1016/S2213-2600(14)70097-9

74. Famous KR, Delucchi K, Ware LB, Kangelaris KN, Liu KD, Thompson BT, et al. Acute respiratory distress syndrome subphenotypes respond differently to randomized fluid management strategy. *Am J Respir Crit Care Med*. (2017) 195:331–8. doi: 10.1164/rccm.201603-0645OC

75. Calfee CS, Delucchi KL, Sinha P, Matthay MA, Hackett J, Shankar-Hari M, et al. Acute respiratory distress syndrome subphenotypes and differential response to simvastatin: secondary analysis of a randomised controlled trial. *Lancet Respir Med*. (2018) 6:691–8. doi: 10.1016/s2213-2600(18)30177-2

76. Zhang Z, Ho KM, Hong Y. Machine learning for the prediction of volume responsiveness in patients with oliguric acute kidney injury in critical care. *Crit Care*. (2019) 23:112. doi: 10.1186/s13054-019-2411-z

77. Parreco J, Hidalgo A, Parks JJ, Kozol R, Rattan R. Using artificial intelligence to predict prolonged mechanical ventilation and tracheostomy placement. *J Surg Res*. (2018) 228:179–87. doi: 10.1016/j.jss.2018.03.028

78. Fabregat A, Magret M, Ferré JA, Vernet A, Guasch N, Rodríguez A, et al. A machine learning decision-making tool for extubation in intensive care unit patients. *Comput Methods Prog Biomed*. (2021) 200:105869. doi: 10.1016/j.cmpb.2020.105869

79. Hur S, Min JY, Yoo J, Kim K, Chung CR, Dykes PC, et al. Development and validation of unplanned Extubation prediction models using intensive care unit data: retrospective, comparative, machine learning study. *J Med Internet Res*. (2021) 23:e23508. doi: 10.2196/23508

80. Chen H, Ma Y, Hong N, Wang H, Su L, Liu C, et al. Early warning of citric acid overdose and timely adjustment of regional citrate anticoagulation based on machine learning methods. *BMC Med Inform Decis Mak*. (2021) 21:126. doi: 10.1186/s12911-021-01489-8

81. Su L, Liu C, Li D, He J, Zheng F, Jiang H, et al. Toward optimal heparin dosing by comparing multiple machine learning methods: retrospective study. *JMIR Med Inform*. (2020) 8:e17648. doi: 10.2196/17648

82. Li D, Gao J, Hong N, Wang H, Su L, Liu C, et al. A clinical prediction model to predict heparin treatment outcomes and provide dosage recommendations: development and validation study. *J Med Internet Res*. (2021) 23:e27118. doi: 10.2196/27118

83. Maviglia R, Michi T, Passaro D, Raggi V, Bocci MG, Piervincenzi E, et al. Machine learning and antibiotic management. *Antibiotics (Basel)*. (2022) 11:304. doi: 10.3390/antibiotics11030304

84. Wang Y, Zhang H, Fan Y, Ying P, Li J, Xie C, et al. Propofol anesthesia depth monitoring based on self-attention and residual structure convolutional neural network. *Comput Math Methods Med*. (2022) 2022:8501948–13. doi: 10.1155/2022/8501948

85. van de Sande D, van Genderen ME, Huiskens J, et al. Moving from bytes to bedside: a systematic review on the use of artificial intelligence in the intensive care unit[J]. *Intensive Care Med*. (2021) 47:750–760. doi: 10.1007/s00134-021-06446-7

86. Barchitta M, Maugeri A, Favara G, Riela P, Gallo G, Mura I, et al. Early prediction of seven-day mortality in intensive care unit using a machine learning model: results from the SPIN-UTI project. *J Clin Med*. (2021) 10:992. doi: 10.3390/jcm10050992

87. Li K, Shi Q, Liu S, Xie Y, Liu J. Predicting in-hospital mortality in ICU patients with sepsis using gradient boosting decision tree. *Medicine*. (2021) 100:e25813. doi: 10.1097/md.0000000000025813

88. Garcia-Gallo JE, Fonseca-Ruiz NJ, Celi LA, Duitama-Muñoz JF. A machine learning-based model for 1-year mortality prediction in patients admitted to an intensive care unit with a diagnosis of sepsis. *Med Intensiva (Engl Ed)*. (2020) 44:160–70. doi: 10.1016/j.medin.2018.07.016

89. Nie X, Cai Y, Liu J, Liu X, Zhao J, Yang Z, et al. Mortality prediction in cerebral hemorrhage patients using machine learning algorithms in intensive care units. *Front Neurol*. (2020) 11:610531. doi: 10.3389/fneur.2020.610531

90. Halonen K, Leppäniemi A, Lundin J, Puolakkainen PA, Kemppainen EA, Haapiainen RK. Predicting fatal outcome in the early phase of severe acute pancreatitis by using novel prognostic models. *Pancreatology*. (2003) 3:309–15. doi: 10.1159/000071769

91. Ding N, Guo C, Li C, Zhou Y, Chai X. An artificial neural networks model for early predicting in-hospital mortality in acute pancreatitis in MIMIC-III. *Biomed Res Int*. (2021) 2021:6638919–8. doi: 10.1155/2021/6638919

92. Niven DJ, McCormick TJ, Straus SE, Hemmelgarn BR, Jeffs L, Barnes TRM, et al. Reproducibility of clinical research in critical care: a scoping review. *BMC Med*. (2018) 16:26. doi: 10.1186/s12916-018-1018-6

93. Weiss CH. Why do we fail to deliver evidence-based practice in critical care medicine? *Curr Opin Crit Care*. (2017) 23:400–5. doi: 10.1097/MCC.0000000000000436

94. Rosa RG, Teixeira C, Sjoding M. Novel approaches to facilitate the implementation of guidelines in the ICU. *J Crit Care*. (2020) 60:1–5. doi: 10.1016/j.jcrc.2020.07.014

95. Liu VX, Morehouse JW, Marelich GP, Soule J, Russell T, Skeath M, et al. Multicenter implementation of a treatment bundle for patients with Sepsis and intermediate lactate values. *Am J Respir Crit Care Med*. (2016) 193:1264–70. doi: 10.1164/rccm.201507-1489OC

96. Eslami S, Abu-Hanna A, Schultz MJ, de Jonge E, de Keizer NF. Evaluation of consulting and critiquing decision support systems: effect on adherence to a lower tidal volume mechanical ventilation strategy. *J Crit Care*. (2012) 27:425.e1. doi: 10.1016/j.jcrc.2011.07.082

97. Trogrlic Z, van der Jagt M, Lingsma H, Gommers D, Ponssen HH, Schoonderbeek JFJ, et al. Improved guideline adherence and reduced brain dysfunction after a multicenter multifaceted implementation of ICU delirium guidelines in 3,930 patients. *Crit Care Med*. (2019) 47:419–27. doi: 10.1097/CCM.0000000000003596

98. Bourdeaux C, Ghosh E, Atallah L, Palanisamy K, Patel P, Thomas M, et al. Impact of a computerized decision support tool deployed in two intensive care units on acute kidney injury progression and guideline compliance: a prospective observational study. *Crit Care*. (2020) 24:656. doi: 10.1186/s13054-020-03343-1

99. Bone RC, Balk RA, Cerra FB. American College of Chest Physicians/Society of Critical Care Medicine consensus conference: definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis. *Crit Care Med*. (1992) 20:864–74. doi: 10.1097/00003246-199206000-00025

100. Levy MM, Fink MP, Marshall JC, Abraham E, Angus D, Cook D, et al. SCCM/ESICM/ACCP/ATS/SIS international Sepsis definitions conference. *Crit Care Med*. (2001) 31:1250–6. doi: 10.1097/01.Ccm.0000050454.01978.3b

101. Singer M, Deutschman CS, Seymour CW, Shankar-Hari M, Annane D, Bauer M, et al. The third international consensus definitions for Sepsis and septic shock (Sepsis-3). *JAMA*. (2016) 315:801–10. doi: 10.1001/jama.2016.0287

102. (中华人民共和国卫生部, 电子病历基本规范(试行)(basic specifications for electronic medical records (trial implementation). (2010).

103. 北京市市场监督管理局 重症医学数据集·患者数据(Intensive Care Medical Dataset Patient Data) DB11/T 1866-2021. 2021.

104. 全国信息技术标准化技术委员会, 工业物联网—数据采集结构化描述规范 *(National Information Technology Standardization Technical Committee, industrial internet of things - structured description specification for data collection)*. (2020).

105. 全国通信标准化技术委员会, 智慧城市 数据融合 第3部分:数据采集规范 *(National Communications Standardization Technical Committee, Smart City data fusion part 3: Data collection specification)*. (2022).

106. 全国信息技术标准化技术委员会大数据标准工作组, 大数据标准化白皮书 (2020版)(big data standards working Group of the National Information Technology Standardization Technical Committee, *White Paper on Big Data Standardization*) (2020).

107. 朱志勇 李, 大数据技术在医疗急重症领域的应用(application of big data technology in the field of medical emergency and critical care). 邮电设计技术 (2016).

108. 家市场监督管理总局与中国国家标准化委员会, 信息技术—数据质量评价指标*GB/T36344-2018 ICS 35.24.01. (state Administration for Market Regulation and China National Standardization Commission, information technology - data quality evaluation indicators)*. (2018).

109. 美国国家标准技术研究院, *NIST*大数据互操作框架:第4册 安全与隐私 *NIST big data interoperability framework: Volume 4, Security and Privacy*. (2013).

110. 中国人民共和国国家质量监督检验检疫总局与中国国家标准化管理委员会. 信息安全技术—大数据服务安全能力要求(information security technology - security capability requirements for big data services). *GB/T 35274-2017*. (2016)

111. 国家市场监督管理总局与国家标准化管理委员会, 信息安全技术—政务信息共享数据信息安全技术要求 *(requirements for government information sharing data information security technology)*. GB/T 39477-2020. (2020).

112. Rieke N, Hancox J, Li W, Milletarì F, Roth HR, Albarqouni S, et al. The future of digital health with federated learning. *NPJ Digit Med*. (2020) 3:119. doi: 10.1038/s41746-020-00323-1

113. Mahmoudi E, Kamdar N, Kim N, Gonzales G, Singh K, Waljee AK. Use of electronic medical records in development and validation of risk prediction models of hospital readmission: systematic review. *BMJ*. (2020) 369:m958. doi: 10.1136/bmj.m958

114. Liang H, Tsui BY, Ni H, Valentim CCS, Baxter SL, Liu G, et al. Evaluation and accurate diagnoses of pediatric diseases using artificial intelligence. *Nat Med*. (2019) 25:433–8. doi: 10.1038/s41591-018-0335-9

115. Rusin CG, Acosta SI, Vu EL, Ahmed M, Brady KM, Penny DJ. Automated prediction of cardiorespiratory deterioration in patients with single ventricle. *J Am Coll Cardiol*. (2021) 77:3184–92. doi: 10.1016/j.jacc.2021.04.072

116. Hamilton PS, Tompkins WJ. Quantitative investigation of QRS detection rules using the MIT/BIH arrhythmia database. *IEEE Trans Biomed Eng*. (1986) BME-33:1157–65. doi: 10.1109/tbme.1986.325695

117. Narula G, Haeberlin M, Balsiger J, Strässle C, Imbach LL, Keller E. Detection of EEG burst-suppression in neurocritical care patients using an unsupervised machine learning algorithm. *Clin Neurophysiol*. (2021) 132:2485–92. doi: 10.1016/j.clinph.2021.07.018

118. Maddali MV, Churpek M, Pham T, Rezoagli E, Zhuo H, Zhao W, et al. Validation and utility of ARDS subphenotypes identified by machine-learning models using clinical data: an observational, multicohort, retrospective analysis. Lancet. *Respir Med*. (2022) 10:367–77. doi: 10.1016/s2213-2600(21)00461-6

119. Walsh SLF, Calandriello L, Silva M, Sverzellati N. Deep learning for classifying fibrotic lung disease on high-resolution computed tomography: a case-cohort study. The lancet. *Respir Med*. (2018) 6:837–45. doi: 10.1016/s2213-2600(18)30286-8

120. Niazi MKK, Parwani AV, Gurcan MN. Digital pathology and artificial intelligence. *Lancet Oncol*. (2019) 20:e253–61. doi: 10.1016/s1470-2045(19)30154-8

121. Massion PP, Antic S, Ather S, Arteta C, Brabec J, Chen H, et al. Assessing the accuracy of a deep learning method to risk stratify indeterminate pulmonary nodules. *Am J Respir Crit Care Med*. (2020) 202:241–9. doi: 10.1164/rccm.201903-0505OC

122. Cheng N, Ren Y, Zhou J, Zhang Y, Wang D, Zhang X, et al. Deep learning-based classification of hepatocellular nodular lesions on whole-slide histopathologic images. *Gastroenterology*. (2022) 162:1948–1961.e7. doi: 10.1053/j.gastro.2022.02.025

123. Walsh SLF, Humphries SM, Wells AU, Brown KK. Imaging research in fibrotic lung disease; applying deep learning to unsolved problems. *Lancet Respir Med*. (2020) 8:1144–53. doi: 10.1016/S2213-2600(20)30003-5

124. Winslow CJ, Edelson DP, Churpek MM, Taneja M, Shah NS, Datta A, et al. The impact of a machine learning early warning score on hospital mortality: a multicenter clinical intervention trial. *Crit Care Med*. (2022) 50:1339–47. doi: 10.1097/CCM.0000000000005492

125. Yuan S, Sun Y, Xiao X, Long Y, He H. Using machine learning algorithms to predict Candidaemia in ICU patients with new-onset systemic inflammatory response syndrome. *Front Med*. (2021) 8:720926. doi: 10.3389/fmed.2021.720926

126. Dastider AG, Sadik F, Fattah SA. An integrated autoencoder-based hybrid CNN-LSTM model for COVID-19 severity prediction from lung ultrasound. *Comput Biol Med*. (2021) 132:104296. doi: 10.1016/j.compbiomed.2021.104296

127. Tiwari P, Colborn KL, Smith DE, Xing F, Ghosh D, Rosenberg MA. Assessment of a machine learning model applied to harmonized electronic health record data for the prediction of incident atrial fibrillation. *JAMA Netw Open*. (2020) 3:e1919396. doi: 10.1001/jamanetworkopen.2019.19396

128. Papp L, Spielvogel CP, Grubmüller B, Grahovac M, Krajnc D, Ecsedi B, et al. Supervised machine learning enables non-invasive lesion characterization in primary prostate cancer with [68Ga]Ga-PSMA-11 PET/MRI. *Eur J Nucl Med Mol Imaging*. (2021) 48:1795–805. doi: 10.1007/s00259-020-05140-y

129. Liang W, Yao J, Chen A, Lv Q, Zanin M, Liu J, et al. Early triage of critically ill COVID-19 patients using deep learning. *Nat Commun*. (2020) 11:1–7. doi: 10.1038/s41467-020-17280-8

130. Su L, Xu Z, Chang F, Ma Y, Liu S, Jiang H, et al. Early prediction of mortality, severity, and length of stay in the intensive care unit of Sepsis patients based on Sepsis 3.0 by machine learning models. *Front Med*. (2021) 8:664966. doi: 10.3389/fmed.2021.664966

131. Su L, Zhang Z, Zheng F, Pan P, Hong N, Liu C, et al. Five novel clinical phenotypes for critically ill patients with mechanical ventilation in intensive care units: a retrospective and multi database study. *Respir Res*. (2020) 21:325. doi: 10.1186/s12931-020-01588-6

132. Liu S, Su L, Liu X, Zhang X, Chen Z, Liu C, et al. Recognizing blood pressure patterns in sedated critically ill patients on mechanical ventilation by spectral clustering. *Ann Transl Med*. (2021) 9:1404. doi: 10.21037/atm-21-2806

133. Landi I, Glicksberg BS, Lee H-C, Cherng S, Landi G, Danieletto M, et al. Deep representation learning of electronic health records to unlock patient stratification at scale. *NPJ Digit Med*. (2020) 3:96. doi: 10.1038/s41746-020-0301-z

134. Hyun S, Kaewprag P, Cooper C, Hixon B, Moffatt-Bruce S. Exploration of critical care data by using unsupervised machine learning. *Comput Methods Prog Biomed*. (2020) 194:105507. doi: 10.1016/j.cmpb.2020.105507

135. Greco M, Caruso PF, Cecconi M. Artificial intelligence in the intensive care unit. *Semin Respir Crit Care Med*. (2021) 42:002–9. doi: 10.1055/s-0040-1719037

136. Richens JG, Lee CM, Johri S. Improving the accuracy of medical diagnosis with causal machine learning. *Nat Commun*. (2020) 11:1–9. doi: 10.1038/s41467-020-17419-7

137. Wei T, Feng F, Chen J, Wu Z, Yi J, He X. Model-agnostic counterfactual reasoning for eliminating popularity bias in recommender system In *Proceedings of the 27th ACM SIGKDD conference on Knowledge Discovery & Data Mining* (2021)

138. Goudet O, Kalainathan D, Caillou P, Guyon I, Lopez-Paz D, Sebag M. Learning functional causal models with generative neural networks In *Explainable and interpretable models in computer vision and machine learning*: Springer (2018). 39–80.

139. Athey S, Tibshirani J, Wager S. Generalized random forests. *Ann Stat*. (2019) 47:1148–78. doi: 10.1214/18-AOS1709

140. Tan J, Xu S, Ge Y, Li Y, Chen X, Zhang Y. Counterfactual explainable recommendation In *Proceedings of the 30th ACM international conference on Information & Knowledge Management* (2021)

141. Tharwat A. Classification assessment methods. *Appl Comput Informat*. (2021) 17:168–92. doi: 10.1016/j.aci.2018.08.003

142. Hong ZZLPY. Predictive analytics with ensemble modeling in laparoscopic surgery: a technical note. *Laparoscopic, Endoscopic and Robotic Surgery*. (2022) 5:25–34. doi: 10.1016/j.lers.2021.12.003

143. Zhang Z, Beck MW, Winkler DA, Huang B, Sibanda W, Goyal H, et al. Opening the black box of neural networks: methods for interpreting neural network models in clinical applications. *Ann Trans Med*. (2018) 6:216. doi: 10.21037/atm.2018.05.32

144. Shimabukuro DW, Barton CW, Feldman MD, Mataraso SJ, das R. Effect of a machine learning-based severe sepsis prediction algorithm on patient survival and hospital length of stay: a randomised clinical trial. *BMJ Open Respir Res*. (2017) 4:e000234. doi: 10.1136/bmjresp-2017-000234

145. Semler MW, Weavind L, Hooper MH, Rice TW, Gowda SS, Nadas A, et al. An electronic tool for the evaluation and treatment of Sepsis in the ICU: a randomized controlled trial. *Crit Care Med*. (2015) 43:1595–602. doi: 10.1097/CCM.0000000000001020

146. Wilson FP, Martin M, Yamamoto Y, Partridge C, Moreira E, Arora T, et al. Electronic health record alerts for acute kidney injury: multicenter, randomized clinical trial. *BMJ*. (2021) 372:m4786. doi: 10.1136/bmj.m4786

147. Selby NM, Casula A, Lamming L, Stoves J, Samarasinghe Y, Lewington AJ, et al. An organizational-level program of intervention for AKI: a pragmatic stepped wedge cluster randomized trial. *J Am Soc Nephrol*. (2019) 30:505–15. doi: 10.1681/ASN.2018090886

148. Wu Y, Chen Y, Li S, Dong W, Liang H, Deng M, et al. Value of electronic alerts for acute kidney injury in high-risk wards: a pilot randomized controlled trial. *Int Urol Nephrol*. (2018) 50:1483–8. doi: 10.1007/s11255-018-1836-7

149. Wilson FP, Shashaty M, Testani J, Aqeel I, Borovskiy Y, Ellenberg SS, et al. Automated, electronic alerts for acute kidney injury: a single-blind, parallel-group, randomised controlled trial. *Lancet*. (2015) 385:1966–74. doi: 10.1016/S0140-6736(15)60266-5

150. Escobar GJ, Liu VX, Schuler A, Lawson B, Greene JD, Kipnis P. Automated identification of adults at risk for in-hospital clinical deterioration. *N Engl J Med*. (2020) 383:1951–60. doi: 10.1056/NEJMsa2001090

151. 齐霜, 毛智, 胡新, et al., 基于专科信息系统建立的重症医学数据库:大型三甲医院重症医学数据库的模式 *(a critical care medicine database established based on specialized information systems: A model of critical care medicine database in large third class hospitals)*. 中华危重病急救医学: p. 743–749.

152. 张素珍,唐素娟,戎珊等, 基于机器学习的重症监护病房脓毒性休克患者早期发生急性肾损伤风险的预测模型构建 *(construction of a machine learning based predictive model for the risk of early acute kidney injury in patients with septic shock in intensive care units)*. 中华危重病急救医学. (2022) 34:255–9.

153. Flechet M, Falini S, Bonetti C, Güiza F, Schetz M, van den Berghe G, et al. Machine learning versus physicians' prediction of acute kidney injury in critically ill adults: a prospective evaluation of the AKIpredictor. *Crit Care*. (2019) 23:282. doi: 10.1186/s13054-019-2563-x

154. Thorsen-Meyer H-C, Nielsen AB, Nielsen AP, Kaas-Hansen BS, Toft P, Schierbeck J, et al. Dynamic and explainable machine learning prediction of mortality in patients in the intensive care unit: a retrospective study of high-frequency data in electronic patient records. *Lancet Digit Health*. (2020) 2:e179–91. doi: 10.1016/S2589-7500(20)30018-2

155. Gutierrez G. Artificial intelligence in the intensive care unit. *Crit Care*. (2020) 24:101. doi: 10.1186/s13054-020-2785-y

156. Sjoding MW, Hofer TP, Co I, Courey A, Cooke CR, Iwashyna TJ. Interobserver reliability of the Berlin ARDS definition and strategies to improve the reliability of ARDS diagnosis. *Chest*. (2018) 153:361–7. doi: 10.1016/j.chest.2017.11.037

157. Fleuren LM, Thoral P, Shillan D, Ercole A, Elbers PWGRight Data Right Now Collaborators. Machine learning in intensive care medicine: ready for take-off? *Intensive Care Med*. (2020) 46:1486–8. doi: 10.1007/s00134-020-06045-y

158. Hravnak M, Pellathy T, Chen L, Dubrawski A, Wertz A, Clermont G, et al. A call to alarms: current state and future directions in the battle against alarm fatigue. *J Electrocardiol*. (2018) 51:S44–8. doi: 10.1016/j.jelectrocard.2018.07.024

159. Yoon JH, Jeanselme V, Dubrawski A, Hravnak M, Pinsky MR, Clermont G. Prediction of hypotension events with physiologic vital sign signatures in the intensive care unit. *Critical Care London, England*. (2020) 24:661. doi: 10.1186/s13054-020-03379-3

Glossary

AES	Advanced encryption standard		LGBM	Light gradient boosting machine
AI	Artificial intelligence		LIME	Locally interpretable agnostic modeling
AKI	Acute kidney injury		LR	Logistic regression
ANN	Artificial neural network		MAE	Mean absolute error
APACHE-II	Acute physiology and chronic health evaluation-II		MIMIC	Medical information mart for intensive care
API	Application programming interface		MOD	Multiple organ dysfunction
ARDS	Acute respiratory distress syndrome		MSE	Mean squared error
AUC	Area under curve		NLP	Natural language processing
AUPRC	Area under precision recall curve		PACS	Picture archiving and communication systems
AUROC	Area under receiver operating characteristic curve		PCA	Principal component analysis
BCT	Big-data clinical trial		PIC	Pediatric intensive care
CATE	Conditional average treatment effect		PMV	Prolonged mechanical ventilation
CDSS	Clinical decision support system		RCT	Randomized controlled trial
CNN	Convolutional neural network		RF	Random forest
COVID-19	Corona virus disease 2019		RMSE	Root mean squared error
CT	Computed tomography		RNA	Ribonucleic acid
DVT	Deep vein thrombosis		RNN	Recurrent neural network
ECG	Electrocardiograph		RR	Risk ratio
EEG	Electroencephalogram		RSA	Ron Rivest, Adi Shamir and Leonard Adleman Algorithm
EHR	Electronic health records		SAP	Severe acute pancreatitis
eICU-CRD	eICU collaborative research database		SAPS	Simplified acute physiology scores
EIT	Electrical impedance tomography		SIRS	Systemic inflammatory response syndrome
EMR	Electronic medical records		SMOTE	Synthetic minority over-sampling technique
ENN	Edited nearest neighbors		SOFA	Sequential organ failure assessment
GBDT	Gradient boosting decision tree		SSC	Surviving sepsis campaign
GRF	Generalized random forest		SVM	Support vector machine
ICH	Intracerebral hemorrhage		THE	Heterogenous treatment effects
ICTCLAS	Institute of Computing Technology, Chinese Lexical Analysis System		THULAC	Thu lexical analyzer for Chinese
ICU	Intensive care unit		TLS	Transport layer security technology
KNN	K-nearest neighbor		UMLS	Unified medical language system
LASSO-GLM	Least absolute shrinkage and selection operator-generalized linear models		USA	United States of America