**ABSTRACT**

**Introduction** Postoperative infection (PI) is one of the main severe complications after cardiovascular surgery. Therefore, antibiotics are routinely used during the first 48 hours after cardiovascular surgery. However, there is no effective method for early diagnosis of infection after cardiovascular surgery, particularly, to determine whether postoperative patients need to prolong the use of antibiotics after the first 48 hours. In this study, we aim to develop and validate a diagnostic model to help identify whether a patient has been infected after surgery and guide the appropriate use of antibiotics.

**Methods and analysis** In this prospective study, we will develop and validate a diagnostic model to determine whether the patient has a bacterial infection within 48 hours after cardiovascular surgery. Baseline data will be collected through the electronic medical record system. A total of 2700 participants will be recruited (n=2000 for model development, n=700 for validation). The primary outcome of the study is the newly PI during the first 48 hours after cardiovascular surgery. Logistic regression penalised with elastic net regularisation will be used for model development and bootstrap and k-fold cross-validation aggregation will be performed for internal validation. The derived model will be also externally validated in patients who are continuously included in another time period (N=700). We will evaluate the calibration and differentiation performance of the model by Hosmer-Lemeshow good of fit test and the area under the curve, respectively. We will report sensitivity, specificity, positive predictive value and negative predictive value in the validation data-set, with a target of 80% sensitivity.

**Ethics and dissemination** Ethical approval was obtained from Medical Ethics Committee of Affiliated Nanjing Drum Tower Hospital, Nanjing University Medical College (2020-249-01).

**Trial registration number** Chinese Clinical Trial Register (www.chictr.org.cn, ChiCTR2000038762); Pre-results.

**INTRODUCTION**

Postoperative infection (PI) is a severe complication after cardiovascular surgery, which can significantly increase mortality and hospitalisation time. Although cardiac surgery techniques and perioperative management have made great progress in recent years, the incidence of infection after surgery has not been significantly reduced. In particular, for low-income countries, the incidence of infection after cardiac surgery is still greater than 10% and some studies believe that infection is responsible for 17% of patient deaths after cardiac surgery.

Administration of surgical antimicrobial prophylaxis is an effective method to reduce the risk of PI. Previous studies indicate that taking antibiotics 30 min before surgery and during the first 48 hours after surgery can effectively reduce the risk of infection in patients after cardiovascular surgery. However, further prolonging the use of preventive antibiotics did not improve...
The prognosis of patients and may be harmful. Therefore, the diagnosis of infection within the first 48 hours after cardiovascular surgery will help guide the use of antibiotics. Unfortunately, to our best knowledge, there is no effective method for early diagnosis of infection after cardiovascular surgery.

Due to the application of cardiopulmonary bypass (CPB), most patients undergoing open heart surgery will inevitably develop a certain degree of systemic inflammatory response syndrome (SIRS). The previous literature has shown that traditional inflammation biomarkers such as leucocytes, neutrophils, and C reactive protein (CRP) are unable to effectively distinguish between SIRS and newly acquired PI. At the same time, some literatures believed that procalcitonin (PCT) had certain value in diagnosing bacterial infections after cardiovascular surgery, but its sensitivity and specificity are poor, especially in the first 48 hours after surgery.

Clinical prediction models have been applied to predict the risk of pneumonia in patients with cardiovascular disease after surgery. However, there is still a lack of well-designed prospective studies to develop and validate clinical diagnostic models related to new infections after open heart surgery. Therefore, we aim to develop a diagnostic model to identify the new infection occurred in the first 48 hours after cardiovascular surgery. It will be able to determine whether the patients are infected after cardiovascular surgery and also help clinicians choose antibiotics more specifically.

OBJECTIVES
In this study, we hope to construct a diagnostic model that can be used to diagnose PI within the first 48 hours after cardiovascular surgery and to guide the use of antibiotics.

METHODS AND ANALYSES
Source of data
The design, conduct and reporting of this study will follow the transparent reporting of a multivariable prediction model for individual prognosis or diagnosis checklists. This work is to develop and validate a diagnostic model based on a prospective study. We will continuously enrol all patients who meet the screening criteria in the Department of Cardiothoracic Surgery of Nanjing Drum Tower Hospital for the development (October 2020 to January 2022) and validation (February 2022 to June 2022) of the diagnostic model.

Participants
Eligibility criteria
1. Inclusion criterion: Adult patients (between 18 and 80 years old) undergoing open cardiovascular surgery.
2. Exclusion criteria: (1) Preoperative body temperature ≥38°C; (2) patients undergoing cardiovascular surgery for trauma, infective endocarditis, neoplasms and malignant tumours; (3) patients who are diagnosed with any other bacterial infectious diseases (such as pneumonia, sepsis); (4) patients who are diagnosed with inflammatory immune diseases and connective tissue diseases; (5) pregnant or lactating women and (6) patients with missing clinical data due to perioperative death or other reasons.

Patients will be required to provide written informed consent for this research. Informed consent will require the use of clinical data, imaging data and serological data during the patient’s hospitalisation.

Outcome
The primary outcome of the study is PI during the first 48 hours after cardiovascular surgery.

Reference diagnostic criteria
Patients will undergo the following aetiological tests within 48 hours after surgery. Meeting either of these criteria is considered to be a confirmed PI.

Postoperative pneumonia
Clinical strategy
The presence of a new or progressive radiographic infiltrate plus at least two of three clinical features (fever greater than 38°C, leucocytosis or leucopenia and purulent secretions).

Bacteriologic strategy
Sputum culture was performed at least twice with fiberoptic bronchoscope or alveolar lavage fluid. Pathogenic bacteria were detected in sputum culture and the first positive result appeared within 48 hours after surgery.

Postoperative pneumonia will only be diagnosed when the patient meets both clinical and bacteriological strategies.

Sepsis
In the first 48 hours after surgery, blood cultures are routinely performed every morning before antibiotic treatment; blood culture will also be performed when the patient’s body temperature is higher than 38.5°C.

Blood cultures sampled from two different sites showed pathogenic infections and will be regarded as PI.

Deep surgical site infection
Culture of the deep secretions of the incision within 48 hours after surgery showed pathogenic infection, which will be regarded as PI.

Diagnostic factors
The factors identified in the previous literature as being potentially indicators of PI and identified from surveys of surgeons and physiotherapists will be measured. These factors are described below.

Clinical factors
Each patient will undergo a standard clinical assessment at admission. The recorded clinical variables will be used for model development, including age, gender, height and weight to obtain body mass index, basal body...
temperature, blood pressure, smoking status, drinking status, hypertension, diabetes, chronic lung disease, atrial fibrillation, chronic renal insufficiency, liver insufficiency, previous surgery and medicine being taken. We will also record the details of the patient's surgeries and complications, as well as the use of various instruments, including surgical procedures, surgical approaches (minimal invasive cardiac surgery, robotic, sternotomy, video assisted), CPB time, aortic cross clamp time, deep hypothermia circulatory arrest, intraoperative blood transfusion volume and type, the amount and type of intraoperative drugs, continuous renal replacement therapy use, intra-aortic balloon pump use, extracorporeal membrane oxygenation use and intraventricular assist device. We will also record the patient's postoperative drainage, body temperature, length of intensive care unit stays, getting out of bed, voluntary sputum expectoration and other key information. Clinical data will be comprehensively collected in the form of case report forms, with a view to assessing the risk factors of infection in patients after cardiovascular surgery in detail.

Serological variables
Serological biomarkers are important for the diagnosis of PI. Although in previous studies, leucocytes, neutrophils and CRP did not show ideal diagnostic value.7 We will further evaluate their prediction ability. We will also include some variables that may be valuable as described below: PCT, albumin, blood glucose, creatine, troponin T/I, interleukin 6, myohaemoglobin and antistreptolysin. During the study, we will continue to explore new potential valuable serological indicators in order to develop a better model. We will test these serological indicators for the first time within 24 hours after the patient is admitted to the hospital and continuously monitor these indicators from the first to the fifth day after surgery.

Radiological variables
During the first to third days after surgery, the patient will have a bedside chest radiograph at least once a day to observe the patient's lung leakage. We will set up a chest radiograph review team composed of two trained doctors to review these chest radiographs separately and classify them into three levels according to the severity of the liquid leakage. Level 1 includes chest radiographs with no obvious exudation, level 2 includes chest radiographs with spot or sheet exudation and level 3 includes large exudates fused into pieces. When two researchers disagree on the results of imaging examinations, imaging diagnostic experts will be consulted to solve any discrepancies. Every non-emergency patient will undergo echocardiography after admission and before discharge. We will evaluate the impact of the patient's left ventricular ejection fraction, ventricular/atrial volume and ventricular wall dyskinesia on the risk of infection after surgery.

Sample size
The calculation of sample size is a vital part of the process of predictive model development and validation. Regarding this question, a recent study has provided a good answer.19 The pmsampsize R package was used. In order to calculate the sample size, we need to make the following settings in the software. First, the primary outcome of our study is binary (being infected or not). According to retrospective analysis, the incidence of PI in our centre is 5%-6% and this is also consistent with previous research reports.1 7 20 Therefore, we assume that the incidence of PI during the study period was 5%. Second, we expect to include 10 candidate variables. For an outcome proportion of 5%, the max(R2 cs) value is 0.33. We conservatively assume that the new model will explain 15% of the variability and the expected R2 cs value is 0.15×0.33=0.05. Then, we used the following parameters in the pmsampsize package: type='b', R2=0.05, parameters=10, prevalence=0.05. The outputs show that the sample size needed to develop a diagnostic model is 1750 and it is expected that 88 events and an events per candidate predictor parameter are 8.75. In order to further improve the reliability of the model, we eventually plan to include 2000 patients. Another 700 patients will be included for model external validation. Therefore, the entire model development and validation process will include a total of 2700 patients.

Missing data
When a variable has more than 5% missing data, we will exclude it from the main analysis. For complete case analysis, missing data will be multiply imputed under a missing at random assumption. We will use multiple imputations, relying on five replications and a chained equation approach method in the R multiple imputation procedure, to account for missing data.

Statistical analysis methods
The development of the diagnostic model will be based on multivariable logistic regression and the calibration and differentiation performance of the model will be evaluated by Hosmer-Lemeshow good of fit test and the area under the curve (AUC), respectively. We will report sensitivity, specificity, positive predictive value and negative predictive value in the validation dataset, with a target of 80% sensitivity.

Software
Statistical analysis of baseline data, sample size calculation, model development and internal and external validation will be completed using R (V.3.5.0).

Development of the diagnostic model
Continuous variables will be expressed as mean±SD or as median (IQR) and compared using Student’s t test or Mann-Whitney U test as appropriate. Normality will be tested by the Shapiro-Wilk test. Categorical data will be compared using the χ2 test or Fisher’s exact test. To explore the influence of each candidate factor on
infection, multivariable logistic regression models will be fitted and OR with their 95% CIs for each candidate factor will be reported. Multivariable analysis will initially include all candidate factors. Reduced multivariable analyses will be considered, if necessary, to examine robustness of the prediction model. Based on a full multivariate regression analysis, the project selection of the model will include those factors that are statistically significant (p<0.05) associated with the outcomes and those that are considered to be clinically important for retention (regardless of statistical significance). The regression model with included factors will be fitted to the cohort data to obtain a final set of parameter estimates to form the model.

Validation of the prediction model
We will perform two internal verification steps to estimate the degree of model overfitting. First, we will divide the patient data used to develop the model into 10 equal parts and use 10-fold cross-validation. We will repeat this operation 10 times to use each subset of the data. Each time, the model will be retrained on 9/10th of the data and then the accuracy statistics will be verified on the remaining 1/10th. This technique provides an average estimate of overfitting and k=10 has been shown to balance the concerns of variance and bias in internal validation. We will also perform a bootstrap internal validation across 500 samples drawn randomly with replacement, which tends to provide an estimate of optimism with lower variance compared with cross-validation. The range of AUCs will be reported for k-fold cross validation and a 95% CIs for the bootstrap internal validation. We will use the data of patients admitted to our centre later in time to construct an external validation set (n=700, not used in the development process). To conduct the external validation, the regression coefficients will be used to score the external validation set. This will make it possible to predict the expected outcome probability and combine it with the observed results, so that the discriminative, calibration and max-rescaled Brier score can be measured in the external validation set.

Risk groups
Due to the serious consequences of infection after cardiovascular surgery, we tend to pay more attention to the sensitivity of the diagnostic model.

DISCUSSION
A number of prediction models have been developed to predict PI.14 21–24 However, all these models used only preoperative or intraoperative information. Different from these previous predictive models to our best knowledge, this is the first study aiming to develop and validate a diagnostic model for PI after cardiac surgery using both preoperative and intraoperative and postoperative data. The variables that will be considered in this diagnostic model include not only demographic data, preoperative factors and surgery-related factors but also postoperative factors.1 4–6 25–31 Therefore, we hope that the model we will construct could provide a more accurate diagnosis of patients who are infected.

Moreover, we will consider as many important perioperative factors as possible, some of which are missed in previous studies. For example, Kilic et al's study mentioned that their prediction model lacked key data such as surgical incision type and ejection fraction.14

Limitation
This is a single-centre study, our development and validation data will be from the Department of Cardiothoracic Surgery of Nanjing Drum Tower Hospital. Therefore, although we have used a variety of methods to validate the model, we still lack data from different hospitals and different regions in China. Second, in order to improve the feasibility of the diagnostic model, we used serological and imaging indicators commonly used in clinical practice. Some potentially effective indicators reported in the previous literature, such as CD64 or pancreatic stone protein, are not included in the evaluation system because they are less clinically used.12 32

ETHICS AND DISSEMINATION
Patients, or their relatives when patients could not consent, will provide written informed consent to participate in the study. Ethical approval was obtained from Medical Ethics Committee of Affiliated Nanjing Drum Tower Hospital, Nanjing University Medical College (2020-249-01). All acquired data will be deidentified, stored electronically and password protected. Data generated in the research will be disseminated via peer-reviewed publications and conference presentations.

Author affiliations
1Department of Cardio-Thoracic Surgery, Nanjing Drum Tower Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Peking Union Medical College Graduate School, Nanjing, China
2Statistical Genetics, QIMR Berghofer Medical Research Institute, Herston, Queensland, Australia
3School of Medicine, The University of Queensland, St Lucia, Queensland, Australia
4Medical Research & Biometrics Center, National Center for Cardiovascular Diseases, Fuwai Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Xicheng District, Beijing, China
5Department of Cardio-Thoracic Surgery, Nanjing Drum Tower Hospital, Nanjing Medical University, Nanjing, Jiangsu, China
6Nanjing Drum Tower Hospital, Xuzhou Medical University, Xuzhou, Jiangsu, China
7Department of Cardio-Thoracic Surgery, Nanjing Drum Tower Hospital, the Affiliated Hospital of Nanjing University Medical School, Nanjing, China

Contributors D-JW and TP have led on design and are overseeing data analysis plans. Data management will be performed by H-TZ, Z-SL, ZC, KP and KZ. Data quality checks will be performed by H-TZ and HZ. Authors X-KH and C-SW are the study statisticians. H-TZ drafted the manuscript. The final dataset will be curated by D-JW, TP and H-TZ, access will be at the discretion of study investigators.

Funding This work was supported by Jiangsu Provincial Key Medical Discipline of the Project of Invigoration Health Care through Science, Technology and Education, Grant number ZDXKA2016019.

Competing interests None declared.
Patient consent for publication  Not required.

Provenance and peer review  Not commissioned; externally peer reviewed.

Open access  This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iD  Dong-Jin Wang http://orcid.org/0000-0001-6499-1972

REFERENCES