Pathways and cost-effectiveness of routine lung cancer inpatient care in rural Anhui, China: a retrospective cohort study protocol

XingRong Shen,1 MengJie Diao,1 ManMan Lu,1 Rui Feng,2 PanPan Zhang,1 Tao Jiang,1 DeBin Wang1

ABSTRACT

Introduction Routine inpatient care (RIC) for patients with cancer forms various pathways of clinical procedures. Although most individual procedures comprising the pathways have been tested via clinical trials, little is known about the collective cost and effectiveness of the pathways as a whole. This study aims at exploring RIC pathways for patients with lung cancer from rural Anhui, China, and their determinants and economic impacts.

Methods and analysis The study adopts a retrospective cohort design and proceeds in five steps. Step 1 defines the four main categories of study variables, including clinical procedures, direct cost and effectiveness of procedures, and factors affecting use of these procedures and their cost and effectiveness. Step 2 selects a cohort of 5000 patients with lung cancer diagnosed between 1 July 2015 and 30 June 2016 from rural Anhui by clustered random sampling. Step 3 retrieves the records of all the inpatient care episodes due to lung cancer and extracts data about RIC procedures, proximate variables (eg, Karnofsky Performance Status, Lung Function Score) of patient outcomes and related factors (eg, stage of cancer, age, gender), by two independent clinician researchers using a web-based form. Step 4 estimates the direct cost of each of the RIC procedures using micro-costing and collects data about ultimate patient outcomes (survival and progression-free survival) through a follow-up survey of patients and/or their close relatives. Step 5 analyses the data collected and explores pathways of RIC procedures and their relations with patient outcomes, costs, cost:effect ratios, and a whole range of clinical and sociodemographic factors using multivariate regression and path models.

Ethics and dissemination The study protocol has been approved by an authorised ethics committee of Anhui Medical University (reference number: 20170312). Findings from the study will be disseminated through conventional academic routes such as peer-reviewed publications and presentations at regional, national and international conferences.

Trial registration number ISRCTN25595562.

INTRODUCTION

Lung cancer has been the most common cancer in the world for several decades.1 The estimated new cases of the disease was 1.8 million in 2012 (12.9% of all cancers), 58% of which occurred in less developed regions. It was also the most common cause of death from cancer worldwide, being responsible for nearly one in five (1.59 million in absolute number) of the total.2 In China, lung cancer incidence has displayed a slight decreasing trend in the past few years, particularly for men. However, it is still the first cancer for men and second for women, accounting for 25.2% of all new cancer cases and 29.5% of all cancer deaths in 2012.3

Routine inpatient care (RIC) for lung cancer consists of a combination of procedures. Patients with possible lung cancer need a detailed history and physical examination first. Then they should undergo posterior-anterior and lateral chest radiographs as well as CT scans of the chest and abdomen. To further confirm and determine the stage and histology of the lesion, other diagnostic methods needed include whole-body fluorodeoxyglucose positron emission tomography, endoscopic ultrasound, sputum cytology, fine needle aspiration, bronchoscopy and others. Following diagnosis of lung cancer, the patients proceed with combined modality
therapies depending on the stage of the disease and comorbidities and complications. Historically, surgery provides the best chance for cure for patients whose lung cancers are limited to the hemithorax and can be totally encompassed by excision.\textsuperscript{4,5} Surgery has been generally used in combination with external beam radiotherapy for control of the primary tumour and regional lymphatics.\textsuperscript{5} In addition, chemotherapy has also been advocated as an integral part of combined modality approaches to earlier stages of the disease.\textsuperscript{7,8} For unselected advanced non-small cell lung cancer, platinum-based combinations have become the standard of care, while cisplatin-based or carboplatin-based doublets are standard for patients with stage IV disease.\textsuperscript{9,10} More recently, epidermal growth factor receptor (EGFR) tyrosine kinase inhibitors have been introduced as second-line and third-line treatments of advanced disease and as first-line treatment for selected patients.\textsuperscript{11}

Given the complex procedures, ensuring quality RIC for patients with lung cancer has been most challenging, and guidelines are widely used in addressing this challenge. Numerous studies have documented positive relations between compliance with guidelines and patient outcomes.\textsuperscript{12,13} However, researchers have also raised concerns about guidelines. One of such concerns refers to the lack of adequate consideration of costs. Most clinical procedures affect disease outcomes and incur considerable costs.\textsuperscript{14,15} Yet guidelines are based on trials focused primarily on effectiveness (eg, survival), with little attention being paid to economic consequences.\textsuperscript{16} Another concern relates to incompatible population between clinical trials and RIC. Clinical trials on which guidelines are based use highly selective populations, while RIC serves a general lung cancer population with different age, performance status and comorbidities.\textsuperscript{17,18} The third concern revolves around uncertain interactions between procedures. Although most individual guideline-recommended procedures (GRPs) have established evidence, they are not used in isolation but in conjunction with others, forming various clinical combinations. Efforts systematically assessing and comparing these combinations are scarce.\textsuperscript{19-22} The fourth concern originates from varied compliance with guidelines since RIC often deviates substantially from guidelines.\textsuperscript{23,24} The cost-effectiveness of these ‘substandard’ or mixed combinations of procedures (partly from guidelines, partly from experiences of individual clinicians) falls far from well-understood.\textsuperscript{25} These all point to a clear need for evaluating RIC even though guidelines are widely available.

All the above concerns are most pertinent to China. First, China has a unique ‘dual’ medical care system in which patients often receive Western medicine and traditional Chinese medicine (TCM) simultaneously or in turn.\textsuperscript{26} Second, China lacks coordinated referral and follow-up mechanisms, and patients with cancer often move freely from one hospital to another for different rounds of inpatient cancer care.\textsuperscript{27} This makes it hard for clinicians to leverage different inpatient care episodes at different time points and hospitals into continuous and synergetic service. Third, China has strong sociocultural norms and financial incentives that hinder cost control and guideline compliance.\textsuperscript{28}

**Study aims**

This study aims at identifying the main pathways of RIC procedures for patients with lung cancer from rural Anhui, China, and exploring determinants of the pathways and their economic impacts. Specific questions to be addressed include the following: (1) What combinations of diagnosis and treatment procedures (or pathways for short) can an individual patient experience during all his or her hospitalisation episodes due to lung cancer-related problems? (2) Which are the most and least frequent pathways? (3) What determines the flow among these pathways? (4) Which are the most and least cost-effective pathways in relation to the other pathways? (5) What factors are associated with the relative cost-effectiveness?

The above ‘pathways’ of inpatient care mean combinations of diagnosis and treatment procedures an individual patient may experience during all his or her hospitalisation episodes due to lung cancer-related problems. Suppose a patient with lung cancer experienced six times/rounds of hospitalised care, and during each of these hospitalisation episodes the patient underwent several diagnosis and treatment procedures, all these procedures form the ‘pathway’ of this particular patient. It is worth noting that findings of the cost-effectiveness analysis are exploratory rather than implying that they are of sufficient robustness to be used to inform policy changes.

**METHODOLOGY**

**Identification of procedures**

The study uses a self-designed and web-based data extraction form to identify major clinical procedures described in any RIC record under concern. The form lists all major RIC procedures under two main domains, that is, diagnostic procedures (eg, chest X-ray, chest CT, neck ultrasonography) and treatment procedures (eg, surgical therapy, chemotherapy, psychobehavioural intervention).

**Estimation of costs**

The study estimates the overall and categorical costs (direct costs only) for each of the RIC procedures (eg, lung function examination, CT, white blood cell count) identified above using micro-costing techniques.\textsuperscript{29,30} Taking the example of lung function examination, categorical costs include costs on personnel, equipment, materials, reagents and others needed in completing the examination, while overall cost of the procedure equals the sum of all these categorical costs. In addition, the study also calculates the grand total cost on individual inpatient by adding up the overall costs on all the clinical procedures he or she has received.
Measurement of effectiveness

The study uses both proximal variables of outcomes (PV) and ultimate outcome (UO) measures of effectiveness of RIC procedures. The UO indicators derive from a follow-up survey about 2.5 years after the first hospitalisation, and include overall survival (OS), progression-free survival (PFS), quality of life (QoL) and quality-adjusted life years (QALYs). Here, QoL is assessed using the widely recognised EuroQol-5 dimensions-3 level (EQ-5D-5L) instrument.31

The PV measures come from RIC records and include Eastern Cooperative Oncology Group, Karnofsky Performance Status, and compiled scores of (1) symptoms (eg, chronic cough, chest pain, wasting syndrome), (2) lung functions (eg, forced vital capacity, forced expiratory volume in 1 s) and (3) image findings (eg, number of nodules identified in the lung, size of the largest nodules, presence of pleura or pericardial effusion). Each of these domain-specific PV scores equals the weighted sum of all subindicators within the domain. For example, the compiled score of ‘lung functions’ equals the sum of the weighted values of forced vital capacity, forced expiratory volume in 1 s and so on. Here the weights come from the coefficients of multivariate regression modelling using an UO indicator (eg, OS) as the dependent variable, while forced vital capacity, forced expiratory volume in 1 s and so on as the independent variables, and the stage of the disease, age, gender and others as the confounding variables.

Calculation of cost-effectiveness

The study adopts the relative cost-effectiveness ratios (RCERs) as the main indicators for measuring cost-effectiveness. Here RCER is defined by the difference in cost between two selected sets of RIC procedures, divided by the difference in their effectiveness. More specifically, RCER=(C/E−C/E), where C and E are the cost and effectiveness in the reference group, and C and E are the cost and effectiveness in the group who has undergone all the procedures in the reference group plus x, a specific procedure under concern.32 Suppose x represents a commonly used TCM that incurs 100 dollars, while r a typical combination of diagnosis and treatment procedures without the TCM. The combination without the TCM costs 1000 dollars and the survival time of patients who have adopted this combination is 1.5 years on average, while the same figure for patients who have used the same combination plus the TCM is 1.51. Then the C =1000, C =1100 dollars and the RCER of the TCM=(1100−1000)/(1.51−1.5)=100 dollars per life year saved. Similarly, RCER=(C/E)/(C/E)=(1100/1.51)/(1000/1.50)=1.09.

Identification of influencing factors

The study also extracts, from RIC records, data about patient factors commonly believed to be linked with disease progression, treatment response and outcomes, and utilisation of RIC procedures. These include (1) sociodemographics (eg, age, gender, body height and weight, education, employment, marital status, medical insurance); (2) risk behaviours and histories (eg, smoking, alcohol drinking, history of cancer among family members); (3) historical and biological test findings (eg, value of anaplastic lymphoma kinase (ALK), V-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog (KRAS), EGFR, programmed death 1 (PD1), carcinoembryonic antigen (CEA), cancer antigen 125 (CA125), pro-gastrin-releasing peptide respectively (proGRP)); and (4) comorbidities and complications (eg, presence of superior vena cava syndrome, brain metastases) and stage of disease. Here, disease staging uses the tumour, node, metastases system, and this staging will be treated as the most important factor throughout the data analysis especially in its effects on the flow of different pathways and their RCER.

Selection of participants

The study is implemented in Anhui, an inland province located in middle and east China. It has a population of 61.4 million, and its per gross domestic product (GDP) and income rank in the middle (the 14th) among all provinces in the nation.33 34 Its social, cultural and economic background is representative of over 80% of the whole population in China.33 34 The province has 68 rural counties, and each of them divides into 10–20 townships. Clustered random sampling was used to select participating counties, townships, patients and RIC case records, and proceeds in five steps. Step 1 classifies all the counties in Anhui into southern, northern and middle areas. Step 2 randomly selects three counties from each of these areas (12 counties in total). Step 3 randomly draws four townships from each of the counties selected (48 townships in total). Step 4 searches the provincial reimbursement database of the New Rural Cooperative Medical System (NRCMS) and identifies all the patients within the selected townships who had been first diagnosed with primary lung cancer between 1 July 2015 and 30 June 2016. Step 5 searches the database again for all episodes of hospitalisation due to lung cancer for patients identified in step 4. NRCMS covers 98% of the rural residents, and the estimated number of patients and admission episodes is about 5000 and 25,000, respectively.

The above sample size was determined by our study purpose of building multivariate models of factors affecting the flow among and the RCER of specific RIC pathways. Patients with lung cancer generally receive four to six rounds of inpatient care. Given the various diagnostic and treatment procedures available, there are hundreds of potential RIC pathways (combinations of diagnosis and treatment procedures from the first to the last round of RIC). We plan to group these pathways into manageable (around 20) categories depending on the resultant distribution of the actual pathways, and we aim to enter 20–30 factors into the multivariate models for each of these categorical pathways. Based on these preconditions and that the sample size of a multivariable
model should generally be 10 times the number of independent variables, we need 250 patients for each pathway. This translates into 5000 patients in total.

**Data collection**

The study obtains data through follow-up survey and data extraction. The follow-up survey applies to all the patients with lung cancer identified above. It solicits information on the following: (1) patient’s disease progression (ie, died, alive with or without progression); (2) if died, date of death; and (3) additional admissions due to the lung cancer not included in the above-mentioned NRCMS database. The survey uses a short structured questionnaire. Administration of the questionnaire starts with a telephone interview (of the patient under concern or his or her close relatives for up to five attempts), followed by a face-to-face interview (of the same respondents for up to two attempts) if the telephone contacts have failed. The recruitment strives to reach over 85% rate of participation. The researchers are trained to record reasons of attrition for each of the patients they have lost so as to allow for assessing potential biases. The data extraction applies to records of all the hospital admission episodes identified via the NRCMS database and the follow-up survey. It uses a structured web-based form and extracts data about the clinical procedures, costs, effectiveness and influencing factors described above. Two experienced clinicians on care of lung cancer perform the data extraction. They visit (on one-by-one base) all the relevant hospitals, ask for permission to examine the full records and fill the worksheet independently first, followed by discussions, if applicable, to solve discordances.

**Data analysis**

The data collected above allow a variety of descriptive and multivariate analyses concerning the costs and effectiveness of RIC. The effectiveness analysis comprises all the UO indicators, including PFS, OS, QoL and disability adjusted life years (DALYs). For each of these UO indicators, the analysis will produce (1) estimation of average rates or values with 95% CIs at different time points after first diagnosis by disease stage, PV indicators, RIC pathways, non-hospital care categories, age range and so on; (2) multivariate regression models using similar variables as independent variables; and (3) path models using disease stage, RIC pathways, non-hospital care categories, age range and so on as exogenous, compiled PV indices as direct endogenous, and individual PV indicators as indirect endogenous variables (figure 1A). The area under the receiver operating characteristic curve will be estimated to assess the predictability of models using binary classifier as the dependent variable (eg, models of PFS, OS).

The cost analysis mainly explores (1) Markov models of the mean cost for managing patients with lung cancer (figure 1B); (2) overall and categorical costs on different rounds of hospitalisation by sociodemographic and selected clinical conditions (figure 2); (3) scatter plot of RIC procedures using the occurrence rate and unit cost of individual procedures as the coordinates; and (4) multivariate regression models of overall and selected categorical costs using disease stage, PV indicators, RIC pathways, non-hospital care categories, age range and so on as independent variables.

The cost-effectiveness analysis focuses primarily on constructing a pathway tree to help estimate expected overall and pathway-specific cost-effectiveness, and identify pathways with the highest or lowest RCER. The tree consists of different branches of combinations of RIC procedures starting from the first to the last episode of inpatient care, labelled with estimated costs and possibilities along the pathways and outcomes at the end of the pathways (figure 3). Relevance of the pathway tree is tested.

![Figure 1](https://example.com/figure1.png)

**Figure 1** Schematic structure of sample multivariate models to be built. D, death; E, effectiveness; e, systematic error; IP, index of proximate variables; L_1+, first line treatment; L_2+, second or third line treatment; N, no active treatment; PA or PB, domain A or proximate indicators of effectiveness; R, remission; X, independent variables.
by, for instance, varying the percentage of patient flowing among the different pathways or the costs of major diagnostic and treatment procedures consisting the branches, and then examining changes in the ranking of the pathways in terms of relative cost-effectiveness. The analysis also pays particular attention to identifying as many as comparable pairs of RIC pathways as possible and calculating RCER accordingly in a hope to uncover potential pathways of practice, policy and research implications.

The pathway tree construction will be done using TreeAge, while descriptive and multivariate model analyses using SPSS V.16. Cases with missing data about a specific item will be excluded from the analysis involving the item, and where applicable the statistical null hypothesis will be rejected at the significance level of $\alpha=0.05$.

### Ethics and dissemination

Participation of hospitals, patients and their relatives is voluntary, and written informed consent is required for all participants. Findings from the study will be disseminated through conventional academic routes such as peer-reviewed publications and presentations at regional, national and international conferences.

### DISCUSSION

The study would share the experience of lung cancer care from the rural Chinese perspective. It is an important sharing of knowledge on population-based lung cancer care, since most economic evidence comes from Europe and North America. In China, TCM is used to complement or replace Western medicine. This results in quite different pathways of lung cancer care that have seldom been well explored in published literature. China has a long history of almost no charges being made for clinical consultations, and most patients are used to paying only for medicines, laboratory tests and equipment-based examinations. This forms a perverse financial incentive for clinicians to order more sophisticated examinations and tests and to overprescribing. China’s lack of referral and follow-up mechanisms also merits particular attention. As an individual patient changes from one hospital (say for the first round of treatment) to another (for the second round treatment), he or she may receive different treatment regimens. Discontinued treatment and follow-up may make it hard for clinicians to base their treatment decisions on observed effects.
Perhaps the most noteworthy findings of the current study may be the description of the pathways of RIC procedures and their economic impacts (figure 2). These pathways will provide easily understandable means for estimating and identifying, among others, the following: (1) which pathways or combinations of procedures happen most or least in routine practice during different rounds of hospitalisation for inpatients suffering from lung cancer in rural China; (2) which pathways (from the first to last round of hospitalisation) incur the highest or lowest direct costs; and (3) which pathways result in the best or worst patient outcome in terms of different UO measures. These have important implications for clinical decision-making as well as policy-making.

Another point worth mentioning refers to the links between the domain-specific proximate (PV) indices and the key UO indicators (eg, OS, PFS, QALYs) generated via a large-scale (involving 5000 patients with lung cancer) retrospective cohort study. They provide useful information for clinicians on care of patients with lung cancer in selecting appropriate procedures to achieve optimal collective contributions to UO. At present, although PV indicators are observed routinely, they are presented to clinicians as individual indicators rather than compiled indices. Given the large number of PV indicators involved and the complex relations between RIC procedures and PV indicators and then UO indicators, it is difficult for practising clinicians to make balanced decisions on their personal experiences.

In addition, this study addresses RIC for lung cancer at hospitals in China from a range of meaningful perspectives. The study reinforces the concepts introduced in the landmark studies of Fisher et al and Wennberg and Fisher, which convincingly demonstrated that high quality was not necessarily associated with high cost. Describing inpatient lung cancer care in a view that its value is directly proportional to outcomes and inversely proportional to costs helps in guiding quality improvement by either better outcomes and/or lower costs. The study calculates and compares the collective costs and effectiveness of different RIC pathways as a whole, and thus informs coordinated inpatient care episodes and procedures at

Figure 3. Anticipated ‘procedure-outcome’ tree of inpatient lung cancer care. Cx, the xth combination of clinical procedures; Ox, the xth patient outcome index/indicator; Px, possibility of using the xth combinations of clinical procedures; Tx, the xth round of hospitalisation.
different time points and hospitals. The study enables RCER estimation for specific GRPs using various combinations of real and uncontrolled RIC procedures as the reference, and thus enhances understanding and application of GRPs established through well-controlled studies.

The study also has limitations. The first limitation concerns data reliability. Although the majority of data are extracted from RIC records kept at hospitals, the study uses self-reported data on QoL and inpatient, outpatient and home care. Self-reports are prone to various biases, including recall problems particularly among the elderly, and over-reporting or under-reporting by the respondents for reasons such as perceived expectations from the researchers or fear of potential worries or distress. These biases may be reduced to a minimum in our study by means of interviewer training, use of chronological recall and probing techniques, and crosschecks of findings from patient interviews, health insurance database and hospital records. Importantly, the study uses the EQ-5D-5L to assess QoL. It has already been tested with adequate reliability both internationally and in China. Regarding non-hospitalised care, the study asks only simple questions about what kind of care the patients have experienced and when and for how long. These questions are relatively memorable and easy to answer. The second limitation relates to selective study content. The study considers only inpatient care, while patients may use various self-treatments and outpatient treatments in addition to inpatient care.40 41 Inpatient and non-inpatient treatment may substitute each other to some extent. These may result in underestimation of the effectiveness of RIC procedures. Fortunately, this underestimation may be offset to a large extent by treating non-hospital care as confounders, and the study data to be collected allow this exercise. Third, the study considers only direct costs rather than full costs, taking both direct and indirect costs into consideration. In addition, different hospitals use different equipment, reagents and medicines. Their quality of records may also vary substantially. These raise compatibility concerns in pooling data from different hospitals together and performing aggregate analysis. Finally, readers may raise concerns about the representativeness of inpatients to larger patients with cancer. Hospitalisation rates documented from other countries varied greatly,42 while similar data from China are scarce. Our estimation, using the data set of the last province-wide Household Health Survey of Anhui, of the proportion of patients with lung cancer who had been admitted to hospitals at least once was as high as 89%.43

Acknowledgements
Refinement and implementation of the protocol are led and supported by the Collaboration Center for Cancer Control of Anhui Medical University, Anhui and Luan Center for Diseases Control and Prevention.

Contributors
XS and MD contributed equally to conceiving this project, facilitating protocol and instrument development, and drafting this manuscript. RF, ML, PZ and TJ are core researchers for cost estimation, record extraction, follow-up survey and data analysis, respectively. DW provided expertise for overall design of the study, and revised and finalised the manuscript. All authors have read and approved the final submission.

Funding
Development of the primitive protocol was supported by the Natural Science Foundation of China (grant number 71503009).

Competing interests
None declared.

Patient consent
Obtained.

Ethics approval
The study protocol had been reviewed and approved by the Biomedical Ethics Committee of Anhui Medical University (reference number: 20170312).

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 and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2018. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

REFERENCES


