

ISSN 1881-7815 Online ISSN 1881-7823

BST

BioScience Trends

Volume 12, Number 2
April, 2018



www.biosciencetrends.com

BST

BioScience Trends



ISSN: 1881-7815
Online ISSN: 1881-7823

CODEN: BTIRCZ

Issues/Year: 6

Language: English

Publisher: IACMHR Co., Ltd.

BioScience Trends is one of a series of peer-reviewed journals of the International Research and Cooperation Association for Bio & Socio-Sciences Advancement (IRCA-BSSA) Group and is published bimonthly by the International Advancement Center for Medicine & Health Research Co., Ltd. (IACMHR Co., Ltd.) and supported by the IRCA-BSSA and Shandong University China-Japan Cooperation Center for Drug Discovery & Screening (SDU-DDSC).

BioScience Trends devotes to publishing the latest and most exciting advances in scientific research. Articles cover fields of life science such as biochemistry, molecular biology, clinical research, public health, medical care system, and social science in order to encourage cooperation and exchange among scientists and clinical researchers.

BioScience Trends publishes Original Articles, Brief Reports, Reviews, Policy Forum articles, Case Reports, News, and Letters on all aspects of the field of life science. All contributions should seek to promote international collaboration.

Editorial Board

Editor-in-Chief:

Norihiro KOKUDO
National Center for Global Health and Medicine, Tokyo, Japan

Yasuhiko SUGAWARA
Kumamoto University, Kumamoto, Japan
Ling WANG
Fudan University, Shanghai, China

Co-Editors-in-Chief:

Xue-Tao CAO
Nankai University, Tianjin, China
Rajendra PRASAD
University of Delhi, Delhi, India
Arthur D. RIGGS
Beckman Research Institute of the City of Hope, Duarte, CA, USA

Managing Editor:

Jianjun GAO
Qingdao University, Qingdao, China

Web Editor:

Yu CHEN
The University of Tokyo, Tokyo, Japan

Chief Director & Executive Editor:

Wei TANG
National Center for Global Health and Medicine, Tokyo, Japan

Proofreaders:

Curtis BENTLEY
Roswell, GA, USA
Christopher HOLMES
The University of Tokyo, Tokyo, Japan
Thomas R. LEBON
Los Angeles Trade Technical College, Los Angeles, CA, USA

Senior Editors:

Xunjia CHENG
Fudan University, Shanghai, China
Yoko FUJITA-YAMAGUCHI
Beckman Research Institute of the City of Hope, Duarte, CA, USA
Na HE
Fudan University, Shanghai, China
Kiyoshi KITAMURA
International University of Health and Welfare, Narita, Japan
Misao MATSUSHITA
Tokai University, Hiratsuka, Japan
Ri SHO
Yamagata University, Yamagata, Japan

Editorial Office

Pearl City Koishikawa 603,
2-4-5 Kasuga, Bunkyo-ku, Tokyo 112-0003, Japan
Tel: +81-3-5840-8764 Fax: +81-3-5840-8765
E-mail: office@biosciencetrends.com

BioScience Trends

Editorial and Head Office

Pearl City Koishikawa 603, 2-4-5 Kasuga, Bunkyo-ku,
Tokyo 112-0003, Japan

Tel: +81-3-5840-8764, Fax: +81-3-5840-8765
E-mail: office@biosciencetrends.com
URL: www.biosciencetrends.com

Editorial Board Members

Girdhar G. AGARWAL <i>(Lucknow, India)</i>	De-Fei HONG <i>(Hangzhou, China)</i>	Yutaka MATSUYAMA <i>(Tokyo, Japan)</i>	Sumihito TAMURA <i>(Tokyo, Japan)</i>
Hirotsugu AIGA <i>(Geneva, Switzerland)</i>	De-Xing HOU <i>(Kagoshima, Japan)</i>	Qingyue MENG <i>(Beijing, China)</i>	Puay Hoon TAN <i>(Singapore, Singapore)</i>
Hidechika AKASHI <i>(Tokyo, Japan)</i>	Sheng-Tao HOU <i>(Ottawa, Canada)</i>	Mark MEUTH <i>(Sheffi eld, UK)</i>	Koji TANAKA <i>(Tsu, Japan)</i>
Moazzam ALI <i>(Geneva, Switzerland)</i>	Yong HUANG <i>(Ji'ning, China)</i>	Satoko NAGATA <i>(Tokyo, Japan)</i>	John TERMINI <i>(Duarte, CA, USA)</i>
Ping AO <i>(Shanghai, China)</i>	Hirofumi INAGAKI <i>(Tokyo, Japan)</i>	Munehiro NAKATA <i>(Hiratsuka, Japan)</i>	Usa C. THISYAKORN <i>(Bangkok, Thailand)</i>
Hisao ASAMURA <i>(Tokyo, Japan)</i>	Masamine JIMBA <i>(Tokyo, Japan)</i>	Miho OBA <i>(Odawara, Japan)</i>	Toshifumi TSUKAHARA <i>(Nomi, Japan)</i>
Michael E. BARISH <i>(Duarte, CA, USA)</i>	Chunlin JIN <i>(Shanghai, China)</i>	Fanghua QI <i>(Ji'nan, Shandong)</i>	Kohjiro UEKI <i>(Tokyo, Japan)</i>
Boon-Huat BAY <i>(Singapore, Singapore)</i>	Kimitaka KAGA <i>(Tokyo, Japan)</i>	Xianjun QU <i>(Beijing, China)</i>	Masahiro UMEZAKI <i>(Tokyo, Japan)</i>
Yasumasa BESSHO <i>(Nara, Japan)</i>	Ichiro KAI <i>(Tokyo, Japan)</i>	John J. ROSSI <i>(Duarte, CA, USA)</i>	Junming WANG <i>(Jackson, MS, USA)</i>
Generoso BEVILACQUA <i>(Pisa, Italy)</i>	Kazuhiro KAKIMOTO <i>(Osaka, Japan)</i>	Carlos SAINZ-FERNANDEZ <i>(Santander, Spain)</i>	Xiang-Dong Wang <i>(Boston, MA, USA)</i>
Shiuan CHEN <i>(Duarte, CA, USA)</i>	Kiyoko KAMIBEPPU <i>(Tokyo, Japan)</i>	Yoshihiro SAKAMOTO <i>(Tokyo, Japan)</i>	Hisashi WATANABE <i>(Tokyo, Japan)</i>
Yuan CHEN <i>(Duarte, CA, USA)</i>	Haidong KAN <i>(Shanghai, China)</i>	Erin SATO <i>(Shizuoka, Japan)</i>	Lingzhong XU <i>(Ji'nan, China)</i>
Naoshi DOHMAE <i>(Wako, Japan)</i>	Bok-Luel LEE <i>(Busan, Korea)</i>	Takehito SATO <i>(Isehara, Japan)</i>	Masatake YAMAUCHI <i>(Chiba, Japan)</i>
Zhen FAN <i>(Houston, TX, USA)</i>	Mingjie LI <i>(St. Louis, MO, USA)</i>	Akihito SHIMAZU <i>(Tokyo, Japan)</i>	Aitian YIN <i>(Ji'nan, China)</i>
Ding-Zhi FANG <i>(Chengdu, China)</i>	Shixue LI <i>(Ji'nan, China)</i>	Zhifeng SHAO <i>(Shanghai, China)</i>	George W-C. YIP <i>(Singapore, Singapore)</i>
Xiaobin FENG <i>(Beijing, China)</i>	Ren-Jang LIN <i>(Duarte, CA, USA)</i>	Judith SINGER-SAM <i>(Duarte, CA, USA)</i>	Xue-Jie YU <i>(Galveston, TX, USA)</i>
Yoshiharu FUKUDA <i>(Ube, Japan)</i>	Lianxin LIU <i>(Harbin, China)</i>	Raj K. SINGH <i>(Dehradun, India)</i>	Benny C-Y ZEE <i>(Hong Kong, China)</i>
Rajiv GARG <i>(Lucknow, India)</i>	Xinqi LIU <i>(Tianjin, China)</i>	Peipei SONG <i>(Tokyo, Japan)</i>	Yong ZENG <i>(Chengdu, China)</i>
Ravindra K. GARG <i>(Lucknow, India)</i>	Daru LU <i>(Shanghai, China)</i>	Junko SUGAMA <i>(Kanazawa, Japan)</i>	Xiaomei ZHU <i>(Seattle, WA, USA)</i>
Makoto GOTO <i>(Tokyo, Japan)</i>	Hongzhou LU <i>(Shanghai, China)</i>	Hiroshi TACHIBANA <i>(Isehara, Japan)</i>	<i>(as of April 2018)</i>
Demin HAN <i>(Beijing, China)</i>	Duan MA <i>(Shanghai, China)</i>	Tomoko TAKAMURA <i>(Tokyo, Japan)</i>	
David M. HELFMAN <i>(Daejeon, Korea)</i>	Masatoshi MAKUUCHI <i>(Tokyo, Japan)</i>	Tadatoshi TAKAYAMA <i>(Tokyo, Japan)</i>	
Takahiro HIGASHI <i>(Tokyo, Japan)</i>	Francesco MAROTTA <i>(Milano, Italy)</i>	Shin'ichi TAKEDA <i>(Tokyo, Japan)</i>	

Policy Forum

- 102 - 108 **Development of health technology assessment in China: New challenges.**
Yingyao Chen, Yao He, Xunyouzhi Chi, Yan Wei, Lizheng Shi
- 109 - 115 **Diagnosis-related group (DRG)-based case-mix funding system, a promising alternative for fee for service payment in China.**
Cuirong Zhao, Chao Wang, Chengwu Shen, Qian Wang

Review

- 116 - 125 **Interventions integrating non-communicable disease prevention and reproductive, maternal, newborn, and child health: A systematic review.**
Kimiyo Kikuchi, Rakesh Ayer, Sumiyo Okawa, Mariko Nishikitani, Fumihiko Yokota, Masamine Jimba, Naoki Nakashima

Original Article

- 126 - 131 **The impact of parental migration on injuries among left behind young people aged 10 years to 24 years in Botswana.**
Lesego Selotlegeng
- 132 - 141 **Internal migration and regional differences of population aging: An empirical study of 287 cities in China.**
Rong Chen, Ping Xu, Fen Li, Peipei Song
- 142 - 148 **Intimate partner violence victimization and HIV infection among men who have sex with men in Shanghai, China.**
Ying Liu, Yuyan Zhang, Zhen Ning, Huang Zheng, Yingying Ding, Meiyang Gao, Frank Y. Wong, Na He
- 149 - 156 ***Astragalus polysaccharide* protects diabetic cardiomyopathy by activating NRG1/ErbB pathway.**
Xiao Chang, Kang Lu, Ling Wang, Min Lv, Wenjun Fu
- 157 - 167 **Distinct pattern of Th17/Treg cells in pregnant women with a history of unexplained recurrent spontaneous abortion.**
Jinfeng Qian, Na Zhang, Jing Lin, Caiyan Wang, Xinyao Pan, Lanting Chen, Dajin Li, Ling Wang
- 168 - 176 **Renal protective effect of Paeoniflorin by inhibition of JAK2/STAT3 signaling pathway in diabetic mice.**
Xinyu Li, Yan Wang, Kun Wang, Yonggui Wu
- 177 - 184 **Effects of three forms of local anesthesia on perioperative fentanyl-induced hyperalgesia.**
Lu Chang, Fang Ye, Quehua Luo, Zewen Wang, Yimin Wang, Zhengyuan Xia, Haihua Shu

CONTENTS

(Continued)

- 185 - 192 **Infiltration characteristics and influencing factors of retroperitoneal liposarcoma: Novel evidence for extended surgery and a tumor grading system.**
Zhen Wang, Jianhui Wu, Ang Lv, Chengpeng Li, Zhongwu Li, Min Zhao, Chunyi Hao
- 193 - 200 **A reliable grading system for prediction of hematoma expansion in intracerebral hemorrhage in the basal ganglia.**
Yongwei Huang, Qiang Zhang, Mingfei Yang

Brief Report

- 201 - 207 **A simple and economical method of gas chromatography-mass spectrometry to determine the presence of 6 pesticides in human plasma and its clinical application in patients with acute poisoning.**
Guiyan Yuan, Rui Zhang, Xuwang Chen, Benjie Wang, Ruichen Guo

Communication

- 208 - 210 **Performing laparoscopic surgery – Perspectives of young Chinese hepatobiliary surgeons.**
Xin Zhao, Peipei Song, Yuhua Zhang, Jiwei Huang
- 211 - 214 **Strengthening maternal and child health in China: Lessons from transforming policy proposals into action.**
Xiaoguang Yang, Shenglan Tang, Gavin Yamey, Xu Qian

Guide for Authors

Copyright

Development of health technology assessment in China: New challenges

Yingyao Chen^{1,2,*}, Yao He^{1,2}, Xunyouzhi Chi^{1,2}, Yan Wei^{1,2}, Lizheng Shi³

¹Key Lab of Health Technology Assessment (Fudan University), National Health Commission, Shanghai, China;

²World Health Organization Collaborating Centre for Health Technology Assessment and Management, Shanghai, China;

³Health Systems Analytics Research Center, School of Public Health and Tropical Medicine, Tulane University, New Orleans, LA, USA.

Summary Health technology assessment (HTA) is a field of scientific policy research that adopts multidisciplinary approaches to conduct systematic evaluation of health technologies and inform decision making. Although achievements have been made by HTA activities among academics, providers, and policy makers, development of the field of HTA in China is fragmented and not yet formally integrated in health policy making processes. All stakeholders need to make more efforts to strengthen HTA knowledge translation and facilitate a decision making process that is based on evidence including HTA findings. This article reviews how the field of HTA has developed in China, analyzes what factors have been influencing China's HTA development, and proposes policy recommendations.

Keywords: Health policy, policy making, decision making, knowledge translation

1. Introduction

Health technology assessment (HTA) is one of the critical evidence-driven decision making processes to inform policy and clinical decision making on the introduction and use of health technologies. Health technologies include pharmaceuticals, devices, diagnostics, procedures, and other clinical, public health, and organizational interventions. The multidisciplinary field of HTA addresses the clinical, economic, organizational, social, legal, and ethical impacts of a health technology, considering its specific healthcare context as well as available alternatives (1). Therefore, HTA is a common policy tool that assists decision makers in using health technology appropriately.

Application of health technology, on the one hand, can create positive social and economic impacts by preventing, diagnosing, and treating diseases, improving quality of life, and caring for people's health. On the other

hand, inappropriate use of health technology may induce negative consequences that, at the individual level, create safety issues and health problems that aggravate the burden of disease and, at the macro level, waste resources and worsen patient-provider relationships. For example, when a college student with a rare cancer received an immunotherapy that was aggressively advertised but had insufficient evidence of effectiveness and died subsequently, the government and the public saw the importance of HTA in avoiding such tragedies (2). If HTA evidence had been used in the introduction of the immunotherapy and its on-line promotion as the top treatment, patients such as the college student might have not suffered physically, emotionally, and financially. Caution is thus applied when evaluating and adopting cancer therapies, robotic surgery systems, high throughput sequencing technologies, and other technologies with insufficient clinical evidence. In addition, determining the tradeoff between the clinical benefits and high costs of new technologies also requires HTA, specifically cost-effectiveness analysis and budget impact analysis, since health resources are limited. For instance, the effective yet highly expensive drugs for hepatitis C and cancers. Therefore, cultivating strengths and avoiding weaknesses of health technology are crucial in the complicated process of using, managing,

*Address correspondence to:

Dr. Yingyao Chen, Key Lab of Health Technology Assessment (Fudan University), National Health Commission, 446 Zhaojiabang Road Building 2 Room 1001, Shanghai 20032, China.

E-mail: yychen@shmu.edu.cn

and monitoring technology development and diffusion, and HTA is the best tool to balance the benefit versus harm, demand versus supply, and appropriateness versus inappropriateness of technology.

HTA has formulated its methodology, including identification of health technology to be assessed, determination of the priority of assessment, implementation of assessment with different methods and highlights, and translation of assessment findings into decisions or policies settings (3).

Since China's most recent round of health reform started in 2009, significant achievements have been made by expanding universal health coverage, improving the access to essential health services, and reducing financial burden. However, new challenges to health status and health system sustainability keep emerging. As the Chinese society goes through the epidemiologic transition, non-communicable diseases (NCDs) which need treatment and management with health technology have become increasingly prevalent in both intensity and extensity. Rapid emergence of cancer and cardiovascular disease has been exerting increasing disease burden on the Chinese population. The prevalence of cancer of all kinds has increased by 63% and that of cardiovascular diseases has increased by 35% from 2006 to 2016 (4), which has been translating into rising dependence on health technologies. Aging has been an integral contributor to the epidemiologic transition. The percentage of population aged 65 and above increased from 8% in 2006 to 11% in 2016 (5), and an aging population contributes to NCD incidences and related technology use significantly. Moreover, due to perverse incentives of technology use in China's health financing scheme, issues of improper use or overuse of technology contribute to rising health care costs and out-of-pocket expenditures and potentially threaten population health (6). As a result, growth rate of health expenditure has been five to ten percentage points higher than that of gross domestic product, and the economic impact of NCDs is projected to be tremendous (7). Confronting such looming cost burdens as well as potential deficit, China's health insurance scheme may not be sustainable. Therefore, the Chinese government established the Healthy China 2030 blueprint in 2016 that places health in the center of the national development agenda and outlines strategies to improve health outcomes, the health industry, and the health system (8). HTA and other policy tools will be increasingly important as the government formulates specific strategies to achieve the Healthy China 2030 goals and improve the insurance package and coverage.

Over the course of HTA development in China, there have been some major activities that contribute to progress, but the policy process still needs to incorporate HTA as routine component. This article reviews HTA development in China, analyzes the

factors that influence China's HTA development, and proposes policy recommendations.

2. HTA development in China

Numerous organizations conduct HTA research and provide evidence base for decision making in China, yet no single organization is equivalent to a national archiving center such as the National Institute for Health and Care Excellence (NICE) as a government entity or the Canadian Agency for Drugs and Technologies in Health (CADTH) as an independent nonprofit. HTA organizations in China have different institutional formats, *e.g.*, governmental institutions, university-based centers, consultant companies, and industry-based centers at both the national and local levels. We surveyed some HTA organizations and browsed their websites in a previous study to understand the development of the field in China (Table 1), but responses were sparse, and there was little information of HTA on the websites. Therefore, we analyzed the HTA-related articles that were written by researchers in China from 1984-2016 to illustrate the changes in the academic realm of HTA (9).

China has made great progress in the field of HTA since its introduction in the 1980s (6). According to Chi's research, a literature retrieval with the keywords "health technology assessment" and "medical technology assessment" in the China Knowledge Resource Integrated Database revealed 351 articles until the end of 2016 (9). Chi categorized HTA activities in China into three stages of development based on the publication year and content of the articles.

The first stage of development in China was from 1993-2000 when several HTA agencies or programs were set up in universities and doing several pilot projects to introduce the concept of HTA to China. The milestone was when Vice Minister of the Ministry of Health advocated policy makers to use HTA at the National HTA Conference in 1999. The main integrated model of HTA and policy translation was assessments of assisted reproductive technology and the Gamma Knife. During this stage only 30 research articles on China's HTA activities were published, on average about four articles were published each year, and the keywords that most articles focused on were "health resources" and "pharmacoeconomics".

The second stage was from 2001 to 2010 when existing HTA organizations formed the engine that drove the development and new ones started to emerge. Compared to the first stage, there was more HTA research on specific technologies such as antenatal screening and large medical equipment such as magnetic resonance imaging scanner and da Vinci surgical system. There were also more articles that introduce or analyze how HTA could serve policy making and how knowledge translation would take place in other countries so that China can learn from them. During this period, 150

Table 1. Examples of health technology assessment (HTA) organizations in China, their institutional format, and year of establishment

Organization	Institutional Format	Year of Establishment
China Medicinal Biotech Association	Industry-based	1993*
Key Lab of Health Technology Assessment (Fudan University), NHFPC	University-based	1994**
Division of Evaluation and Translational Research, Development Center for Medical Science and Technology, NHFPC	Governmental institution	1994
Chinese Cochrane Center	University-based	1999
Center for Pharmacoeconomic Evaluation and Research, Fudan University	University-based	2002
Center for Evidence-Based Medicine, Fudan University	University-based	2004
Evidence-based Medicine Center, Lanzhou University	University-based	2005
Division of Health Policy Evaluation and Technology Assessment, China National Health Development Research Center, NHFPC	Governmental institution	2007
Shanghai Health Technology Assessment Research Center	Governmental institution	2011
Division of Technology Assessment, National Center for Medical Service Administration, NHFPC	Governmental institution	2015
Center for Health Policy and Technology Assessment, Peking University	University-based	2017
Research Center for Health Technology Assessment of Hubei Province	University-based	2017

Note: NHFPC: National Health and Family Planning Commission. * Although the China Medicinal Biotech Association was established in 1993, it did not start conducting HTA until 2007. ** The predecessor, Research Center for Medical Technology Assessment, was established in 1994, and then it was designated as the Key Lab of HTA in 2004 by NHFPC. The World Health Organization (WHO) designated the Key Lab as the WHO Collaborating Centre for Health Technology Assessment and Management in 2007.

research articles on China's HTA activities had been published, on average 15 articles were published each year, and the most popular keywords were "health resources," "antenatal examination," and "alert."

The third stage was rapid development of the field of HTA from 2011 until now. Compared to the second stage, there have been more articles on the safety, cost-effectiveness, and ethical implication of specific technologies as well as more studies on HTA methodologies. From 2011-2016, 171 research articles on China's HTA activities had been published, on average about 34 articles were published each year, and the most popular keywords were "health policy," "public health," and "biguanide." Nine HTA research projects received funding in 2017 from the National Natural Science Foundation, the top public organization that supports scientific research in China. Although HTA projects make up an insignificant proportion of the total number of grants, the number of HTA proposals and grants have been increasing in recent years (10).

Recent changes that took place during the third stage at two key ministries governing the health system also demonstrated clear progress in incorporating HTA to policy making. The National Health and Family Planning Commission (NHFPC) has issued policies to guide and strengthen the implementation and use of HTA in China since 2016. The policies are the *Guidelines of Developing Health Science and Technology Innovation* and *Guidelines of Facilitating the Transfer*

and Translation of Health Science and Technology Outcomes. The NHFPC departments that oversee health technology-related issues are going to adopt HTA and to apply it to decision making in general. NHFPC has also started to incorporate HTA into specific policy making processes. For instance, the Department of Drug Policy and Essential Medicine of NHFPC asked medical device manufacturers to participate in the price negotiation of four groups of high-value medical consumables in 2017 by submitting HTA evidence of the products along with other required information (11). Although NHFPC did not mandate that external parties should conduct the HTA studies, the inclusion of HTA indicated that NHFPC took concrete steps to use HTA to inform policy making.

Moreover, China's national drug reimbursement policy, *i.e.*, the National Reimbursement Drug List (NRDL), has included evidence of HTA or pharmacoeconomic evaluation as one of the criteria of determining whether a drug would be covered by the national health insurance schemes. In 2017 the Ministry of Human Resources and Social Security (MOHRSS) updated NRDL with an evidence-based, value-driven approach rather than repeating the past process of making reimbursement for drugs passively without scientific input (12). MOHRSS invited more than 4,000 experts in clinical service delivery, pharmacology, health economics, HTA, and health insurance to vote for the inclusion or exclusion of a preliminary list of drugs. If a drug has high clinical value but is expensive,

patented, or exclusive, it could be included in price negotiation before officially becoming part of NRDL (Figure 1). MOHRSS identified 44 drugs to undergo price negotiation and asked the drug companies to provide evidence that supported inclusion of their drugs in NRDL. Two key sections of the evidence, namely clinical value and economic value, coincided with two of the main areas of HTA. This shows MOHRSS was starting to make decisions based on quantitative evidence rather than relying on qualitative expert consensus. After the evidence-based negotiation process, the prices of 36 drugs were successfully reduced (81.8%), and the drugs

were added to NRDL. The average price reduction was 44%, and the highest reduction was 70%. The drugs consisted of five traditional Chinese medicines and 31 western drugs; 15 drugs were for cancer (41.7%); and while 21 were imported, 15 were domestic (13). This marked a significant milestone for MOHRSS to use HTA in making policies that increase access to effective drugs while maintaining health expenditure at a sustainable level.

The series of policy changes are also catalyzed by researchers' persistent effort of submitting HTA research findings to the government as reference for policy

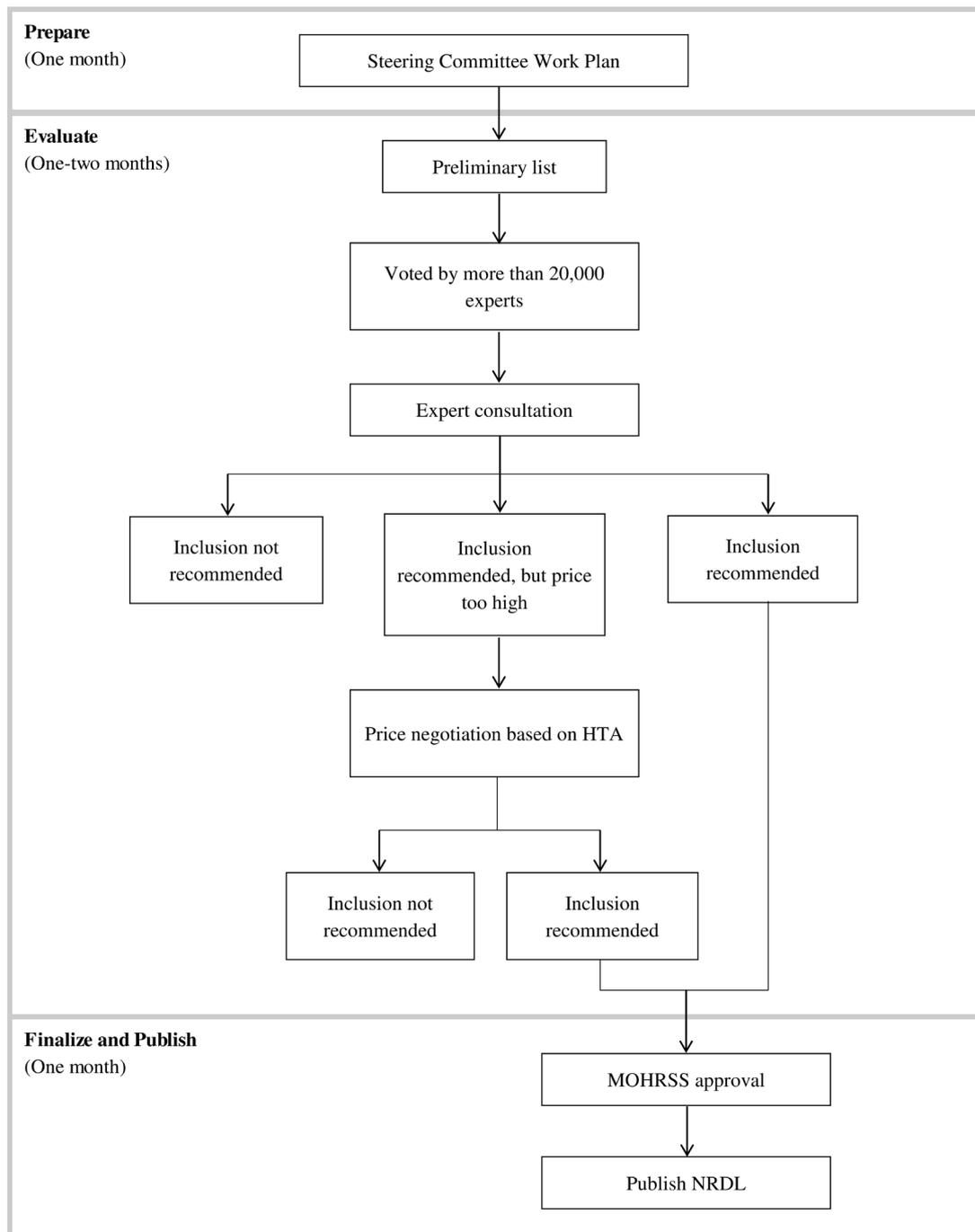


Figure 1. The process of selecting and adding drugs to the National Reimbursement Drug List (NRDL) by China's Ministry of Human Resources and Social Security (MOHRSS) in 2017.

making. While learning HTA techniques continuously, Chinese scholars and institutions have been conducting HTA research, promoting its application, and expanding research teams. In 2017 alone, at least twenty-seven national and regional academic conferences on HTA took place in 12 major cities throughout China, which made HTA one of the most active fields in terms of knowledge exchange and dissemination (14). As a result, abundant research achievements and academic exchanges have been made to serve health policy making.

However, health policy- and decision making in China still largely relies on experience rather than research evidence, since HTA application remains fragmented, sporadic, and not fully embedded into the policy making process as a mandatory component. Chi's research using the HTA mapping instrument designed by Wija Oortwijn *et al.* shows that China's HTA development score is 75 out of 146 (51.4%), which is lower than the scores of developed countries, and one of the least developed domains is HTA implementation in policy and practice (4 out of 10 points) (9,15).

The core issue is thus the lack of HTA integration into the regulatory, reimbursement, and decision making processes, which results in few pathways for HTA findings to translate into policies or decisions. As old technologies evolve and new ones emerge, the need for translational medicine keeps growing, and requirement of precise, scientific decision making becomes more rigorous. Therefore, the field of HTA in China needs further development to strengthen evidence-based decision making and management, increase the efficiency of resource utilization, and improve health outcomes.

3. Factors influencing HTA development in China

Using John Kingdon's three-stream model of the policy process, we analyze factors that influence issues in the problem, policy, and political streams regarding HTA development in China (16).

3.1. Problem Stream

In terms of problems, pharmaceuticals, equipment, and devices still cause safety issues frequently. Effectiveness and affordability of the clinical application of new technologies are sometimes debatable if not unclear. Adjustment of health financing and reimbursement scheme for technology is also largely arbitrary and unreasonable, given little HTA evidence has been taken into account. For example, the proportion of drug costs has been decreased while that of device costs has been increased, even if the appropriateness of individual drugs and devices varies. Market access for innovative technology and the NRDL's slow response to people's needs and technology development are always criticized by the public. As the government simplifies the regulatory process, balancing deregulation of market

approval, pricing of health technology, and assurance of safety, effectiveness, affordability, and social acceptability becomes increasingly critical.

3.2. Policy Stream

In terms of the policy stream, there are three issues, namely HTA has little influence on and integration in health policy in China, the policy process that involves HTA is disorganized and fragmented, and the contribution of HTA to policy making has mediocre quality.

First, implementing HTA to provide policy solutions to issues of health technology remains inadequate in China, and HTA is a policy tool that decision makers seldom recognize and acknowledge. According to the report of the 2015 Global Survey on Health Technology Assessment by National Authorities, almost half of the 111 World Health Organization member countries that participated in the survey (46%) have issued legislations to incorporate HTA findings in decision making; two thirds of the countries, not only developed countries but also middle-income ones such as Brazil and Malaysia, have established national HTA agencies; and over 90% of the countries have adopted HTA guidelines (17). China yet needs to take the aforementioned actions and join the common trend across nations of developing HTA as an essential policy tool. Although HTA has become one of the building blocks of China's recent health policies such as NRDL, the government has not required that reimbursement policy making should be based on HTA by an independent third party. The lack of such a requirement may jeopardize fairness and credibility of the evaluation process that informs policy making.

Second, a national HTA organization has not emerged in China to have the authority of setting and implementing HTA-related priorities and standards as well as coordinating participation in policy making among the numerous existing HTA organizations in China. The lack of such an organization contributes to the fact that HTA has not been integrated to the policy process. The governing scheme that oversees health technology-related issues is also fragmented, since it involves at least four ministries and many ministry departments, which makes role determination, coordination, and communication highly challenging. In addition, policies and decisions of market approval and technology selection are based on expert consensus and past experience rather than detailed technology assessments. As market approval undergoes deregulation, more emphasis has been placed on monitoring health technology after market entry, but long-term mechanisms or laws are absent to ensure effective use of HTA in monitoring.

Third, the number of HTA researchers and agencies is low, the scope of application is narrow, and the quality of reports needs improvement, which deters policy makers

from considering HTA evidence seriously. HTA activities in China are mostly scientific research projects that are sporadic and confined in disciplinary silos. This makes it difficult for the few HTA researchers to participate consistently in health technology management and policy making, implementation, and evaluation. The needs of HTA researchers are not aligned with those of health policy makers, which further hinders communication between the two parties. Moreover, the amount of HTA research in China is still small, and the quality of the research is uneven across different research teams due to the lack of standards and guidelines.

3.3. Political Stream

Analysis of the political stream and social climate regarding HTA development in China shows multiple constraints. Awareness of HTA among decision makers and health care providers is highly limited. Chi's research shows about 60% of policy makers report that they are not capable of reading and understanding HTA findings, and few agree that HTA has played a significant role in health policy making in China (18). Only 3.5% fully understand the concept of HTA, while 48.5% do not know the concept at all (9). Among the ones who understand HTA, many find navigating HTA research papers and deciphering health economic jargons difficult. Unable to comprehend the implication of HTA for solving real healthcare problems, decision makers often regard HTA as theoretic research that does not produce practical results. Therefore, HTA findings are seldom considered as important evidence for decision making or an integral part of the policy process. Furthermore, the general public, which includes users of health technology, does not possess the basic awareness of HTA, making the social climate un conducive to HTA development. Although the use of HTA may influence the health and need for health technology of every member of the society, very few people know the existence of HTA. Without public participation and support, HTA can hardly draw policy makers' attention.

Lack of awareness results in the fact that government leadership has little acknowledgement of HTA and does not value its importance. During its 30 years of development in China, HTA has only been mentioned in several recent policies rather than playing a significant role in shaping the health reform and guidelines of healthcare management. HTA has had few advocates among senior government leaders who exert great influence on policy making, and HTA researchers have had little resource and few effective channels to reach policy makers and promote the use of HTA.

4. Policy recommendation

While problems with the field of HTA in China are increasingly prevalent and pronounced, solutions are

unidentified, and political will remains low. Fortunately, government entities such as NHFPC and MOHRSS have established momentum of using HTA that will affect policy making in at least the next few years. Therefore, China could adopt the following approaches to keep the momentum and apply HTA effectively to meeting the needs of all stakeholders and facilitating evidence-based policy making.

First, China needs an organizing system of health technology assessment and management that promotes implementation and application of HTA and receives external evaluation that holds the government accountable. Exemplars which are worth learning from include the UK NICE, CADTH, and the incorporative system in the US where the government and market compliment and collaborate with each other. China shall adapt successful international experiences to its context and follow the principle that market determines resource allocation so that the government can perform other duties and functions.

An independent consortium of HTA agencies, rather than one single agency, should be established with government funding, and each agency of the consortium would play its unique role while collaborating with each other. The publicly funded independent status of the consortium ensures the objectivity and nonprofit nature of the evaluation research. The consortium helps relevant ministries conduct HTA; suggests HTA findings as reference for decision making; and has the authority to recommend policies and decisions of health technology market approval and selection for public procurement, health insurance, and public health service package. The consortium shall have an HTA expert committee that organizes various ministries and departments to identify key health technologies to assess and coordinate HTA implementation. The consortium shall also promote the development and application of HTA by encouraging all stakeholders of health technology, *e.g.*, health providers, social security that determines health service reimbursement, and manufacturers, to require or provide evidence of external HTA for decision making.

After designating the consortium to implement and develop HTA, China may call for an appraisal mechanism that reviews the health technology reimbursement process and engages stakeholders to critique and discuss HTA reports and recommend the final decision of reimbursement. The appraisal mechanism can only take effect if reviewers are willing to replace isolated observations and experiences with empirical evidence as the basis of policy discussions. HTA reports, the cornerstone of the mechanism, shall synthesize evidence on the safety, effectiveness, affordability, and social implication of a technology to inform the government of technology management. An appraisal committee shall then propose policy recommendations based on thorough discussions of the reports and make the decision. The government agency,

independent consortium, or professional association responsible for HTA shall initiate the appraisal process and form the appraisal committees that consist of health technology researchers, government officials, and interest group representatives.

Several measures should be taken to improve the quality of HTA research in China so that HTA findings will become increasingly convincing and be translated into policy. An official HTA guideline should be established to dictate the standard of quality and uphold the principle of academic integrity and transparency. Moreover, before publication, HTA reports should undergo peer review by domestic and international experts from related disciplines to gain multifaceted feedback. Technology manufacturers that conduct assessment of their products need to disclose the entire process and findings of the research to the government and the public.

Last but not least, active producers and users of HTA in China should take every opportunity to showcase the value of HTA. The policy climate for HTA development has been improving, as new health technologies emerge, population needs increase and diversify, and the health reform progresses and recognizes the value of scientific evidence in general. Therefore, it is significant for all stakeholders to produce meaningful and influential research with practical implications, propel the momentum, and open the policy window for HTA in China.

Acknowledgements

We sincerely thank the China Medical Board Collaborating Programs for supporting the development of the field of HTA in China as well as the Key Lab of HTA (Fudan University), NHFPC continuously.

References

- Health Technology Assessment international. What is HTA? <http://www.htai.org/htai/what-is-hta.html> (accessed January 9, 2018).
- British Broadcasting Corporation. China investigates search engine Baidu after student's death. May 3, 2016. <http://www.bbc.com/news/business-36189252#> (accessed March 7, 2018).
- O'Reilly D, Campbell K, Vanstone M, Bowen JM, Schwartz L, Assasi N, Goeree R. (2015). Evidence-based decision making 3: Health technology assessment. In: Clinical Epidemiology, Methods in Molecular Biology (Parfrey P & Barrett B, eds.). Humana Press, New York, NY, USA, 2015; pp. 1281.
- Institute for Health Metrics and Evaluation. Graph illustration Global Burden of Disease Compare 2016. Vizhub. <https://vizhub.healthdata.org/gbd-compare/> (accessed March 7, 2018).
- National Bureau of Statistics of China. National data. <http://data.stats.gov.cn/english/easyquery.htm?cn=C01> (accessed March 7, 2018).
- Chen Y, Banta D, Tang, Z. Health technology assessment development in China. *Int J Technol Assess Health Care*. 2009; 25:202-209.
- World Bank. Healthy China: Deepening health reform in China. <http://www.worldbank.org/en/country/china/publication/healthy-china-deepening-health-reform-in-china> (accessed March 7, 2018).
- State Council of the People's Republic of China. Healthy China 2030 blueprint. http://www.gov.cn/zhengce/2016-10/25/content_5124174.htm (accessed March 7, 2018; in Chinese).
- Chi, X. Mapping of health technology assessment in China: Current situation analysis and international comparison (Unpublished master's thesis). 2017. Fudan University, Shanghai, China.
- Zuo G. Has the spring of health technology assessment come in 2017? https://mp.weixin.qq.com/s?__biz=MzlyNjE4Mzc5Mg==&mid=2650285603&idx=1&sn=cb50fa909e4177e896ef940793892835&chksm=f078d93cc70f502a7468ebd3176ba61a95737ba05d04d0236f221e27e9f1acfd0bd8dc075219&scene=18#wechat_redirect (accessed March 8, 2018; in Chinese).
- National Health and Family Planning Commission. Announcement of the national price negotiation of high-value medical consumables. <http://www.nhfpc.gov.cn/yaozs/zhgl/201709/505172a3724e4a88a0864e5f5a9b3fc9.shtml> (accessed March 7, 2018; in Chinese).
- Ministry of Human Resources and Social Security. Announcement of the National Reimbursement Drug List of the basic national health insurance, occupational injury insurance, and maternity insurance (2017 edition). http://www.mohrss.gov.cn/yiliaobxs/YILIAOBXSzhengcewenjian/201702/t20170223_266775.html (accessed April 16, 2018; in Chinese).
- Ministry of Human Resources and Social Security. Results of the inclusion negotiation of the National Reimbursement Drug List. http://www.mohrss.gov.cn/SYrlzyhshbzb/dongtaixinwen/buneyaowen/201707/t20170719_274189.html (accessed March 8, 2018; in Chinese).
- Zuo G. At least 27 conferences on health technology assessment and pharmacoeconomics took place in 2017. https://mp.weixin.qq.com/s?__biz=MzlyNjE4Mzc5Mg==&mid=2650285607&idx=2&sn=915c60b495b0eaf4f969a8a8113c875e&chksm=f078d938c70f502e65df40b49cca8a9870268f4d181b868ef7740f10ed0c9f92f8747523e3c3&scene=18#wechat_redirect (accessed March 8, 2018; in Chinese).
- Oortwijn W, Broos P, Vondeling H, Banta D, Todorova L. Mapping of health technology assessment in selected countries. *Int J Technol Assess Health Care*. 2013; 29:424-434.
- Kingdon J. (1984). *Agendas, alternatives, and public policy* (2nd ed.). Harper Collins, New York, NY, USA, 1984.
- World Health Organization. 2015 global survey on health technology assessment by national authorities: Main findings. http://www.who.int/health-technology-assessment/MD_HTA_oct2015_final_web2.pdf (accessed March 7, 2018).
- Shi L, Mao Y, Tang M, Liu W, Guo Z, He L, Chen Y. Health technology assessment in China: Challenges and opportunities. *Global Health Journal*. 2017; 1:11-20.

(Received March 14, 2018; Revised April 23, 2018; Accepted April 25, 2018)

Diagnosis-related group (DRG)-based case-mix funding system, a promising alternative for fee for service payment in China

Cuirong Zhao¹, Chao Wang², Chengwu Shen¹, Qian Wang^{1,*}

¹Department of Pharmacy, Shandong Provincial Hospital Affiliated to Shandong University, Ji'nan, China;

²Department of Rehabilitation Medicine, Shandong Provincial Hospital Affiliated to Shandong University, Ji'nan, China.

Summary

Fee for services (FFS) is the prevailing method of payment in most Chinese public hospitals. Under this retrospective payment system, medical care providers are paid based on medical services and tend to over-treat to maximize their income, thereby contributing to rising medical costs and uncontrollable health expenditures to a large extent. Payment reform needs to be promptly implemented to move to a prospective payment plan. The diagnosis-related group (DRG)-based case-mix payment system, with its superior efficiency and containment of costs, has garnered increased attention and it represents a promising alternative. This article briefly describes the DRG-based case-mix payment system, it comparatively analyzes differences between FFS and case-mix funding systems, and it describes the implementation of DRGs in China. China's social and economic conditions differ across regions, so establishment of a national payment standard will take time and involve difficulties. No single method of provider payment is perfect. Measures to monitor and minimize the negative ethical implications and unintended effects of a DRG-based case-mix payment system are essential to ensuring the lasting social benefits of payment reform in Chinese public hospitals.

Keywords: Fee-for-service, diagnosis-related groups, case-mix funding system, China

1. Introduction

The main sources of funding for most Chinese public hospitals are government allocations, income from fees for medical services and medicines, and other forms of income. Government financing is allocated by the provincial financial bureau in a certain proportion based on hospital beds every year. However, the ratio of government allocations is relatively low, accounting for less than 10% of the sources of funding for most public hospitals. More than 90% of public hospitals' funds come from fees for medical services and medicines, such as performing procedures that require high-tech equipment and dispensing drugs. The fee-for-service (FFS) payment system, with a disease-specific cap for every admission is the prevailing method of payment

in Chinese public hospitals. The retrospective payment system, which reimburses hospitals based on clinic visits, examinations, and treatment programs (1), is feasible and simple to administer. However, improper incentives as part of China's dominant FFS payment model are largely responsible for the rising costs of health care (2,3). Under the Chinese FFS system, the government controls the pricing of medical services, so the prices for advanced care and drugs were set higher than their actual cost while the prices for basic care were set lower (2,4). For hospitals to obtain 90% of their funds, physicians were encouraged to prescribe expensive and profitable medications or diagnostic tests that were not always beneficial to patients (4). Consequently, over-treatment and over-prescription caused by the FFS payment system were widespread in China, leading to rising medical costs (4). Along with limited insurance coverage, the rapid increase in health care costs resulted in public concern that 'it is too difficult to see a doctor and too expensive to seek health care' (5). Payment reform needs to be promptly implemented to move away from a retrospective FFS payment model to a prospective payment plan.

In order to contain the continuing growth of health

Released online in J-STAGE as advance publication April 15, 2018.

*Address correspondence to:

Dr. Qian Wang, Department of Pharmacy, Shandong Provincial Hospital Affiliated to Shandong University, 324 Jing Wu Road, Ji'nan, 250021, China.

E-mail: wangqian.9186@163.com

expenditures, the Chinese Government was called upon by the World Bank to convert its health system from a purely FFS system to a mixed payment system as early as 1997 (6). The mixed payment system may involve methods of prospective payment such as global budgets, capitation, and case-mix-based payment systems. In response, the Government committed to medical reform and it announced its intention to reform hospital payments by moving from a FFS system to a prospective payment system. These methods of prospective payment, such as capitation, global budgets, and a diagnosis-related group (DRG)-based case-mix payment system, are conducive to cost containment and have been piloted in some Chinese cities (7). A DRG-based case-mix payment system is a promising alternative (8). The case-mix funding system, with its superior efficiency and containment of costs, has gradually become the principal means of reimbursing hospitals in many countries (9-11).

2. A case-mix system based on DRGs

2.1. The Case Mix System

The Case Mix System (CMS), a hospital-based decision-making support system, was developed by Providence Hospital (Southfield, Michigan) in conjunction with the consulting and public accounting firm Arthur Andersen & Co. and with developmental and financial support from Blue Cross/Blue Shield of Michigan (12). A case-mix system database is generated from the hospital's patient medical records abstracting system that includes clinical descriptions of treated patients, billing data documenting the treatment rendered, and cost data. The comprehensive administrative database can be used for financial management, planning, analysis, and research. *Via* this database, the case-mix system offers a useful measure for intra-institutional assessment of medical practices and comparison of performance across hospitals (12,13). The case-mix system classifies instances of patient treatment and it reflects the aggregate risk of all patients at a hospital (13). Each instance of patient admission to discharge is referred to as an episode of care. Patient treatment episodes are designed to create classes that include patients with similar clinical characteristics and that possess relatively homogeneous patterns of resource consumption (9). The case-mix system varies in patient condition, disease mix, and the volume of patients treated, and it identifies the financial impact of changes in medical practices (12). In a health care system, development of case-mix classifications is driven by both socio-political and technical factors. The intended scope and use of the classification, the underlying population size, and the quality and depth of the coded data are the three technical factors that influence development of case-mix classifications, including how a case-mix classification is developed

and how many end classes best reflect the complexity of treatment in a hospital (14).

2.2. Diagnosis-related groups

Resource consumption has been widely used as a proxy for the severity of illness. The hospital Case Mix Index (CMI) was developed worldwide to contain costs. The CMI is usually reported in diagnosis-related groups (DRGs) based on International Classification of Diseases (ICD) coding (12,13). Many countries do not have DRG systems, but they do use an ICD-derived CMI that relies on grouping of ICD codes (10). Although DRGs were originally developed solely as a measure of hospital performance by researchers at Yale University in the 1960s, they are now extensively used as hospital payment mechanisms, increasing transparency, improving efficiency and facilitating hospital management in the United States, Europe, Australia, and elsewhere (9,10). DRGs are more widely preferred than capitation in practice (5), and in recent years DRGs have gradually become one of the most remarkable prospective payment systems around the world (11). In principle, DRGs can also be used to reimburse hospitals for acute and non-acute inpatient care, though they are primarily used to provide reimbursement for acute inpatient care (15).

In DRG-based hospital payment systems, the payment categories (hospital services) are defined by DRGs (16). By definition, DRGs are "diagnosis-related" groups of patients that have homogenous patterns of resource consumption and that are clinically meaningful at the same time. Therefore, treated cases that are classified into the same DRG are economically and medically similar (15). Classification of cases in DRGs is based on the following variables: principal and secondary diagnoses, type of treatment, patient age and sex, surgery, the existence of co-morbidities and complications, discharge status, and the procedures performed (11,15,16). Once a patient is discharged and leaves the hospital, case notes generated during the episode of care are examined and assigned a corresponding DRG category according to the ICD. The risk level of DRGs is also ranked on the basis of illness severity and complications. A complex case involving a more severe illness that requires more difficult treatment and interventions is classified into a high-risk DRG (17). Classification is accomplished by improved information systems, including a complete medical records abstracting system that requires physicians to follow uniform standards on writing medical records. Moreover, coding principles should be unified to match the hospital information system (18).

An exhaustive patient case classification system is one core design characteristic of a DRG-based payment system. Another core design characteristic is the payment formula (Figure 1) (15), which is based on the base rate multiplied by a relative cost weight specific to each DRG. The base rate, which can cover all costs

$$\begin{array}{c}
 \boxed{\text{DRG-based}} \\
 \boxed{\text{payment rate for}} \\
 \boxed{\text{a case group}} \\
 \text{(DRG variant)}
 \end{array}
 =
 \begin{array}{c}
 \boxed{\text{Cost weights}} \\
 \text{(expenditure ceiling)}
 \end{array}
 \times
 \begin{array}{c}
 \boxed{\text{Base rate}}
 \end{array}
 \times
 \begin{array}{c}
 \boxed{\text{Adjustment factors}}
 \end{array}$$

Figure 1. Payment formula for a DRG-based payment system.

Table 1. Differences between a fee-for-service (FFS) and a case-mix funding system

Differences	FFS	Case-mix funding system
Charges are assessed based on	Medical services	Case mix
When charges are assessed	After medical services are provided	Before medical services are provided
Incentive mechanism for assessment of charges	Public hospital income is related to the amount and price of medical services	Public hospital income depends on the gap between actual medical expenses incurred by the hospital and assessed charges based on a case mix

or specific costs, is usually a monetary value and is the same for all DRGs. When setting the base rate, cost considerations should be taken into account, and costs are largely influenced by the total funds available. An expenditure ceiling can be created (15). The cost weight, independent from budgetary concerns in principle, is a relative measure that reflects the relative use of resources linked to a specific DRG in comparison to other DRGs. Adjustments must be made to determine the right relative costs for a country. The cost weight may be high for one DRG but low for another DRG. Setting relatively higher cost weights allows overcompensation for highly cost-effective services (15). The DRG-based payment rate needs to be adjusted for numerous reasons such as regional differences and additional funding for teaching hospitals. Adjustment factors are used as a tool to adjust the DRG-based payment rate. Therefore, DRG variants, cost weights, expenditure ceilings, and adjustment factors are the core design components of DRGs, namely payment formulae, which are the basis for payment standards. The establishment of payment standards helps to control medical expenditures, and it also provides compensation for reasonable medical costs and it reasonably benefits medical institutions. This is conducive to increasing staff motivation and hospital development (19).

Under a DRG-based financing system, payment rates are pre-defined for each patient treatment episode in a particular DRG category. Hospitals are paid a standard amount according to the number and type of cases they treat, regardless of the actual cost of caring for an individual (11,20). The use of DRGs, which introduces an element of financial risk for service providers, allows a sophisticated case-mix payment system as a way to reduce health care overuse (6). Accordingly, accurate diagnosis and classification, reasonable payment standards, and other relevant data

play a vital role in a hospital's funding system. To a great extent, they decide the hospital's reimbursement and economic survival (17). In many countries, professional medical associations, consultants, and specialists formally participate in the process of selecting, defining, and updating classification criteria (16). Reasonable payment standards require a scientific pricing mechanism that relies on the market with multiple layers of governance (19).

3. Comparative analysis of fee-for-service and case-mix funding systems

3.1. The key difference between FFS and case-mix funding systems

The key difference between FFS and case-mix funding systems is the mechanism of payment reimbursement. This includes differences in how charges are assessed, when charges are assessed, and the incentive mechanism for assessment of charges (Table 1) (21). Under an FFS payment system, charges are assessed for medical services such as examinations, diagnosis, prescriptions, surgeries, and other forms of care. This means medical services are essentially a form of income for hospitals. Under a case-mix funding system, in contrast, hospitals receive a fixed rate for each admission depending on a patient's DRG category. The pre-defined fee for treating patients in a single DRG category sets a limit on the overall expenses for individual patients, regardless of the actual cost of care (8,11). In essence, medical services are therefore transformed into costs for hospitals. As shown in Table 1, charges are assessed after medical services are provided under an FFS payment system but before medical services are provided under a case-mix funding system. Accordingly, an FFS payment system is a type of retrospective payment system with

a mechanism of supervising the quality of medical treatment. Theoretically, patients could pay for health care depending on its quality, thereby monitoring the quality of medical treatment and prompting its efficiency. In actuality, however, patients have great difficulty monitoring the quality of medical treatment due to the asymmetry of information between the patient and the physician. Patients must accept how charges are assessed. In contrast, a case-mix funding system is a type of prospective payment system with a mechanism of controlling medical costs. The pre-defined fee for a case mix compels hospitals to contain costs when providing medical services. The differences in how charges are assessed and when charges are assessed determines the difference in the incentive mechanism for assessment of charges, which is the essential difference between FFS and case-mix funding systems. The manner in which charges are assessed is essentially the incentive mechanism for public hospitals. Under an FFS payment system, public hospital income is related to the amount and price of medical services, *i.e.*, the higher the amount and price of a medical service, the greater the income. Therefore, an FFS payment system inevitably encourages hospitals to pursue maximum benefits by increasing the quantity and price of those medical services. Under a case-mix funding system, in contrast, public hospital income depends on the gap between the actual medical expenses incurred by a hospital and assessed charges based on a case mix. The pre-defined fee for a case mix according to assessed charges is prepaid to the hospital. If the actual medical expenses are less than the pre-defined fee, the remainder reverts to the hospital. Conversely, if the actual medical expenses are higher than the pre-defined fee, the excess is borne by the hospital. As a consequence, hospitals should increase the efficiency and quantity of medical services and contain costs to make a profit.

3.2. *Advantages and disadvantages of the FFS payment system*

The general consensus appears to be that an FFS payment system is largely responsible for overprovision of care, inefficiency, cost inflation, uncontrollable health expenditures, and even an erosion of medical ethics (2,22). As providers are paid based on medical services under an FFS payment system and there is an asymmetry of information between the patient and the physician, physicians are encouraged to maximize their income by overprescribing drugs and diagnostic tests with high profit margins, prolonging the duration of hospitalization, and delivering unnecessary services (2,6,22). As an example, 75% of patients suffering from a common cold are prescribed antibiotics in China (23). Unnecessary services including overordering expensive drugs and tests in turn lead to runaway cost inflation, wasted resources, poor quality, and unaffordable

health care (3). Moreover, there is an asymmetry of information between national health officials and the physician. This hampers the reasonable pricing of medical services. In fact, substantial differences in how charges were assessed have been noted among hospitals at same level. These weaknesses of the FFS payment system are not conducive to supervision and evaluation of hospitals by national health officials, thereby resulting in high management costs. Moreover, patients are prevented from predicting medical expenses. The increased medical expenses caused by the FFS payment system may result in patient antipathy to medical charges and tension in the doctor-patient relationship.

Although an FFS payment system has the numerous disadvantages described earlier, it also has advantages (summarized in Table 2). Calculation of charges in the FFS payment system is simple and easy to understand. Under an FFS payment system, physicians are compensated for providing the best medical services to patients based on professional standards (22). The FFS payment system can increase the motivation of physicians to a greater extent. Adequate medical care is delivered in a timely manner. Patients with a complex condition will not be rejected by physicians since hospitals will not be bearing a financial risk. Physicians can provide pro bono care to poorer patients who cannot afford it by decreasing or waiving payment. This shortfall in revenue can be balanced out by charging rich patients more (22). Moreover, uncontrolled medical expenses under an FFS payment system are conducive to development of new medical technologies, which always cost more.

3.3. *Advantages and disadvantages of a case-mix funding system*

A DRG-based case-mix funding system is often expected to increase the transparency of hospital performance, contain resource consumption by standardizing reimbursement, and increase efficiency by encouraging appropriate care and discouraging unnecessary care (15,20). Under a DRG-based case-mix payment system, hospitals receive a fixed rate for each admission depending on a patient's diagnosis, regardless of the actual cost of caring for the individual. The limited overall expenses for individual patients represent a higher risk of insolvency for hospitals and increased attention to their bottom line (24), presumably leading to the mitigation of over-treatment under the FFS payment system (8). To make a profit, hospitals are encouraged to control costs, reduce the patient's lengths of stay (LOS), and to simultaneously increase the number of inpatient admissions (11,24). In this manner, a DRG-based case-mix payment system can achieve substantial control of medical costs, effective utilization of health resources, and gains in efficiency (5). A unified payment standard under a DRG-based case-

Table 2. Advantages and disadvantages of an FFS and a case-mix funding system

Payment system	Advantages	Disadvantages
FFS	Calculation of charges is simple and easy to understand.	Largely contributes to over-provision of medical services and rising health care costs.
	Physician motivation can be increased to a greater extent and adequate medical care is delivered in a timely manner.	Medical services are difficult to price and prices differ among hospitals.
	Hospitals bear no financial risk and never shunt patients to other facilities.	Not conducive to supervising and evaluating hospitals and results in higher management costs.
	Conducive to development of new medical technology.	May cause tension in the doctor-patient relationship
Case-mix funding system	Controls health care cost	Difficult to implement and requires support through relevant policies and information technology.
	Constrains resource consumption and increases efficiency	Increased hospital readmission rates
	Increases comparability across hospitals and is conducive to supervising and evaluating hospitals	Provides inadequate medical care
	Relieves tension in the doctor-patient relationship	Rejects patients with a complex condition

mix payment system can increase comparability across hospitals and is conducive to supervision and evaluation of hospitals by national health officials. In addition, the DRG-based financing system greatly helps to improve hospital transparency (20). Once patients are included in the DRG-based payment system, they are notified of medical procedures and the total expense. This should relieve tension in the doctor-patient relationship.

Although considerable evidence from developed countries and several developing countries has indicated that a DRG-based case-mix funding system has advantages in terms of cost containment and improved efficiency (5), this system might also encourage opportunistic practices (summarized in Table 2). To enhance profits, hospitals may classify a patient in a DRG with higher reimbursement, readmit patients, or select low-cost patients and treatments that are more lucrative (24). This payment system has been found to increase hospital readmission rates (23) and increase case volumes (15). Moreover, physicians might provide inadequate medical care and reject patients with a complex condition since their treatment would result in greater resource consumption (2). For this reason, adequate medical care may not be delivered in a timely manner like that under an FFS payment system. Lastly, the complex implementation of a DRG-based case-mix funding system requires strong support from relevant policies and information technology, including hospital information systems, medical record-keeping and diagnostic coding, and clinical pathways. Its application is limited by actual circumstances at a hospital.

4. DRG-based case-mix funding system in China

DRGs are not new to China. As early as 1994, Huang *et al.* analyzed the feasibility of applying the 1990 version

of the All Patient DRGs (AP-DRGs) to hospitals in Beijing (26). They found that the coefficient of variation (CV) for the LOS was 95% in medical DRGs and 73% in surgery and that the CV for costs was 129% in medical DRGs and 94% in surgical DRGs. In 2001, Gong *et al.* (27) launched a study of the feasibility and applicability of introducing Australian refined DRGs (AR-DRGs version 4.0) to care at Chinese hospitals in Chengdu. They found that the Australian refined DRGs provide a good basis for development of Chinese DRGs. Chinese hospitals in relatively developed provinces have started to explore DRG-based methods of payment (28). In 2004, Shanghai experimented with a prospective payment system in which a reimbursement cap was imposed on each DRG (11). A preliminary study determined hospitalization costs for corresponding diseases. Based on those results, the insurance payment standard was set at the average medical insurance cost from several years prior. To facilitate hospital payment reform and performance assessment, Beijing launched a local DRG system (BJ-DRGs) in 2009 (29). The BJ-DRGs consisted of 108 groups, including all types of acute inpatient services, and performed similarly with regard to within-group homogeneity and predictive validity (30). In 2011, Beijing piloted a shift from an FFS payment system or a payment system based on 108 groups to a DRG-based payment system at six tertiary general hospitals (6). That payment reform in Beijing hospitals led to reductions in health expenditures and out-of-pocket payments without any increase in hospital readmission rates or any shift in costs from cases covered by the DRGs to cases not covered by those groups (6).

As information technology has recently advanced in China, medical information systems in most hospitals have been enhanced. At the same time, medical record-

keeping and diagnostic coding have been further standardized, clinical pathways have been developed and standardized, and prices for medical services have become relatively uniform. Moreover, the Chinese Government has paid closer attention to payment reform and enhanced supervision of the quality of medical care (18). All of these factors have greatly promoted the implementation of DRGs. However, the current DRG-based case-mix payment system used to reimburse hospitals for in-patient services in China covers only a few common conditions such as acute appendicitis and hysteromyoma. These conditions are readily diagnosed and their treatment protocols are widely accepted. The most widespread method of payment in China is through incomplete DRGs, which are not real DRGs but a system of paying for individual conditions (11). Due to the level of medical standards and limited resources in China, reasonable grouping of all diseases will be difficult. Thus, a bridge between the FFS payment system and the DRG-based case-mix payment system must be created. Using gastrointestinal diseases (11), hypertension, and coronary heart disease (31), Wang *et al.* showed how to prioritize steps in the implementation of DRGs in the context of limited resources. They found that screening common diseases, studying key factors for medical costs, and simplifying the classification of DRGs will greatly increase the efficiency of payment reform. Nevertheless, China's complex social and economic conditions differ across regions, so establishment of a national payment standard will take time and involve difficulties.

5. Conclusion

In China, the inflationary FFS payment system has been found to create inefficiencies, inflate costs, create waste, and result in unaffordable and poor-quality health care (2,32). China needs to take the right step and reform that payment system. The DRG-based case-mix payment system, in which the financial risk is shifted from payers to providers, is more likely to contain costs, lower the financial barriers to care, and improve efficiency. This prospective method of payment is a promising alternative to the FFS payment system in China. However, no single method of provider payment is perfect. Each has its own advantages and disadvantages and can induce unintended behavior. Methods of payment such as a DRG-based case mix provide strong incentives to reduce costs and probably tend to reduce quality, under-provide care, and even exclude sick patients. Hence, payment reform should be implemented carefully and more rigorous and longitudinal studies should be conducted to verify the superiority and suitability of a DRG based case-mix payment system. The full implementation of a DRG-based case-mix payment system would require additional measures to monitor and minimize its

negative ethical implications and unintended effects (8). Rigorous evidence-based assessment emphasizing the quality of medical care and outcomes and measures to dissociate care providers' profit motives from the incentives of physicians they employ are essential and need to be promptly implemented in China (2). Reforming payment systems at public hospitals essentially means reforming the incentive mechanism (21). The establishment of reasonable payment standards based on a scientific pricing mechanism could bring public hospitals reasonable benefit from medical cost control and prompt them to actively implement payment reform (19). Moreover, professional ethics and norms in medicine should be re-instituted to ensure the lasting social benefits of payment reform in Chinese public hospitals.

References

1. Bai CH, Zhang BY, Zhang Y, Luo L, Hua Y, Wang ZF, Fu HP, Yao Q, Hao M. The policy on "total control and structural adjustment" of medical expenses and changes in hospital payments. *Zhongguo Yi Yuan Guan Li*. 2002; 22:18-20. (in Chinese)
2. Yip WC, Hsiao W, Meng Q, Chen W, Sun X. Realignment of incentives for health-care providers in China. *Lancet*. 2010; 375:1120-1130.
3. Gao C, Xu F, Liu GG. Payment reform and changes in health care in China. *Soc Sci Med*. 2014; 111:10-16.
4. Hui EC. The contemporary healthcare crisis in China and the role of medical professionalism. *J Med Philos*. 2010; 35:477-492.
5. Hu S, Tang S, Liu Y, Zhao Y, Escobar ML, de Ferranti D. Reform of how health care is paid for in China: Challenges and opportunities. *Lancet*. 2008; 372:1846-1853.
6. Jian W, Lu M, Chan KY, Poon AN, Han W, Hu M, Yip W. Payment reform pilot in Beijing hospitals reduced expenditures and out-of-pocket payments per admission. *Health Aff (Millwood)*. 2015; 34:1745-1752.
7. Li C, Yu X, Butler JR, Yiengprugsawan V, Yu M. Moving towards universal health insurance in China: Performance, issues and lessons from Thailand. *Soc Sci Med*. 2011; 73:359-366.
8. Jin P, Biller-Andorno N, Wild V. Ethical implications of case-based payment in China: A systematic analysis. *Dev World Bioeth*. 2015; 15:134-142.
9. Palmer G, Reid B. Evaluation of the performance of diagnosis-related groups and similar casemix systems: Methodological issues. *Health Serv Manage Res*. 2001; 14:71-81.
10. Ammar W, Khalife J, El-Jardali F, Romanos J, Harb H, Hamadeh G, Dimassi H. Hospital accreditation, reimbursement and case mix: Links and insights for contractual systems. *BMC Health Serv Res*. 2013; 13:505.
11. Wang Z, Liu R, Li P, Jiang C, Hao M. How to make diagnosis related groups payment more feasible in developing countries- A case study in Shanghai, China. *Iran J Public Health*. 2014; 43:572-578.
12. Bauer PS, Rinaldo JA. The Case Mix System. *J Med Syst*. 1984; 8:55-63.

13. Richardson D, Tarnow-Mordi WO, Lee SK. Risk adjustment for quality improvement. *Pediatrics*. 1999; 103 (1 Suppl E):255-265.
14. Jackson T, Dimitropoulos V, Madden R, Gillett S. Australian diagnosis related groups: Drivers of complexity adjustment. *Health Policy*. 2015; 119:1433-1441.
15. Mathauer I, Wittenbecher F. Hospital payment systems based on diagnosis-related groups: Experiences in low- and middle-income countries. *Bull World Health Organ*. 2013; 91:746-56A.
16. Quentin W, Scheller-Kreinsen D, Geissler A, Busse R, EuroDRG group. Appendectomy and diagnosis-related groups (DRGs): Patient classification and hospital reimbursement in 11 European countries. *Langenbecks Arch Surg*. 2012; 397:317-326.
17. Müller ML, Bürkle T, Irps S, Roeder N, Prokosch HU. The diagnosis related groups enhanced electronic medical record. *Int J Med Inform*. 2003; 70:221-228.
18. Zhang H, Tian WH. Exploring problems with hospital reimbursement and their solutions using a case mix approach. *Zhongguo Wei Sheng Zhi Liang Guan Li*. 2015; 22:25-28. (in Chinese)
19. Zhao Y, Pan XY. Reflection on the reform of fee payment by disease in public hospitals. *Zhongguo Yi Yuan Guan Li*. 2013; 33:6-9. (in Chinese)
20. Fourie C, Biller-Andorno N, Wild V. Systematically evaluating the impact of diagnosis-related groups (DRGs) on health care delivery: A matrix of ethical implications. *Health Policy*. 2014; 115:157-164.
21. Zhao Y. Rethinking reform of the method of payment in public hospitals. *Zhongguo Wei Sheng Shi Ye Guan Li*. 2016; 11:818-821. (in Chinese)
22. Ikegami N. Fee-for-service payment - An evil practice that must be stamped out? *Int J Health Policy Manag*. 2015; 4:57-59.
23. Yip W, Hsiao WC. The Chinese health system at a crossroads. *Health Aff (Millwood)*. 2008; 27:460-468.
24. Herwartz H, Strumann C. Hospital efficiency under prospective reimbursement schemes: An empirical assessment for the case of Germany. *Eur J Health Econ*. 2014; 15:175-186.
25. Anderson GF, Steinberg EP. Hospital readmissions in the Medicare population. *N Engl J Med*. 1984; 311:1349-1353.
26. Huang H. Feasibility of applying DRGs in Beijing hospital management. *Zhongguo Yi Yuan Guan Li Za Zhi*. 1994; 10(3):131-136. (in Chinese)
27. Gong Z, Duckett SJ, Legge DG, Pei L. Describing Chinese hospital activity with diagnosis related groups (DRGs). A case study in Chengdu. *Health Policy*. 2004; 69:93-100.
28. Liang X, Guo H, Jin C, Peng X, Zhang X. The effect of new cooperative medical scheme on health outcomes and alleviating catastrophic health expenditure in China: A systematic review. *PLoS One*. 2012; 7:e40850.
29. Jian W, Huang Y, Hu M, Zhang X. Performance evaluation of inpatient service in Beijing: A horizontal comparison with risk adjustment based on Diagnosis Related Groups. *BMC Health Serv Res*. 2009; 9:72-79.
30. Jian WY, Lu M, Cui T. Evaluating performance of local casemix system by international comparison: A case study in Beijing, China. *Int J Health Plann Manage*. 2011; 26:471-481.
31. Wang Z, Liu R, Li P, Jiang C. Exploring the transition to DRGs in Developing Countries: A case study in Shanghai, China. *Pak J Med Sci*. 2014; 30:250-255.
32. Mechanic RE, Altman SH. Payment reform options: Episode payment is a good place to start. *Health Affairs*. 2009; 28:w262-271.

(Received November 17, 2017; Revised February 21, 2018; Accepted March 23, 2018)

Interventions integrating non-communicable disease prevention and reproductive, maternal, newborn, and child health: A systematic review

Kimiyo Kikuchi^{1,*}, Rakesh Ayer², Sumiyo Okawa², Mariko Nishikitani¹, Fumihiko Yokota¹, Masamine Jimba², Naoki Nakashima³

¹ Institute of Decision Science for a Sustainable Society, Kyushu University, Fukuoka, Japan;

² Department of Community and Global Health, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan;

³ Medical Information Center, Kyushu University Hospital, Fukuoka, Japan.

Summary

Reproductive, maternal, newborn, and child health (RMNCH) care services could be critical entry points for preventing non-communicable diseases in women and children. In high-income countries, non-communicable diseases screening has been integrated into both the medical and public health systems. To integrate these services in low- and middle-income countries, it is necessary to closely examine its effectiveness and feasibility. In this systematic review, we evaluated the effectiveness of integrating gestational and non-gestational non-communicable diseases interventions and RMNCH care among women and children in low- and middle-income countries. This systematic review included randomized and quasi-randomized controlled trials published from 2000 to 2015. Participants included reproductive-age women, children < 5 years old, and RMNCH care providers. The included interventions comprised packaged care/services that integrated RMNCH services with non-communicable disease care. The outcomes were maternal and/or infant mortality and complications, as well as health care service coverage. We analyzed six studies from 7,949 retrieved articles. Yoga exercise ($p < 0.01$) and nutritional improvements ($p < 0.05$) were effective in reducing gestational hypertension and diabetes. Additionally, integrating cervical cancer and RMNCH services was useful for identifying potential cervical cancer cases. Interventions that integrate non-communicable disease care/screening and RMNCH care may positively impact the health of women and children in low- and middle-income countries. However, as primary evidence is scarce, further research on the effectiveness of integrating non-communicable disease prevention and RMNCH care is warranted. (*Review Registration: PROSPERO International prospective register of systematic reviews (CRD42015023425).*)

Keywords: Systematic review, maternal health, low- and middle-income countries, noncommunicable diseases

1. Introduction

The health of women and children is identified as an international priority issue in the Sustainable Development Goals (1). Women tend to be more vulnerable due to increased risk exposure and

poor access to care, and more than half of all non-communicable disease (NCD)-related deaths were estimated to occur in women (2). Furthermore, among women, the major cause of death is NCDs; these account for nearly 65% of deaths in women (3), with cardiovascular disease being the number one cause of death in this population (4). Some NCDs specifically affect women and adolescent girls, such as breast and cervical cancers (5). These cancers are largely treatable in developed countries, but women in resource-limited settings have fewer treatment options (6-8).

In addition, reproductive-age women have particular

*Address correspondence to:

Dr. Kimiyo Kikuchi, Institute of Decision Science for a Sustainable Society, Kyushu University. 3-1-1, Maidashi, Higashi-ku, Fukuoka, 812-8582, Japan.

E-mail: kikuchi.kimiyo.715@m.kyushu-u.ac.jp

NCD risks (e.g., obesity, hypertension, and cancers), and certain NCDs (e.g., diabetes) are often accelerated during pregnancy and childbirth. Gestational hypertension may affect the mother's womb and lead to eclamptic attack and placental abruption, as well as growth retardation of the fetus. Further, gestational diabetes, which affects one-fourth of pregnant women (2), may cause stillbirth, premature labor, difficult delivery, or neonatal hypoglycemia. Finally, poor maternal health may further result in chronic conditions of their child, such as asthma, heart disease, and diabetes (9).

Reproductive, maternal, newborn, and child health (RMNCH) care services could be critical entry points for NCD prevention among women and children in resource-limited settings (5). These services do not only prevent gestational NCDs but also NCDs throughout the lifetime of the woman. This is because nearly 80% of pregnant women have at least one antenatal visit and most children are vaccinated, while their access to additional health care is largely limited in low- and middle-income countries (10). Typical RMNCH care in these countries includes antenatal care, birth attendance, postnatal care, vaccination for children, and family planning. One previous study revealed that, during the antenatal care, up to 45% of gestational NCDs could be prevented through proper assessment and education (11). Moreover, treating gestational diabetes can reduce the number of newborn deaths (11), and congenital heart disease can be successfully treated if adequate postnatal care is provided (12). In addition, preeclampsia may increase the risk of cardiovascular disease by two to four times (13), and preventing gestational diabetes may prevent the development of type 2 diabetes in the future (14). Therefore, the World Health Organization (WHO), other international health authorities, and academia recommend that healthcare for NCDs should be provided as part of an integrated approach for RMNCH services (2,5,15).

Integrating health care services is considered an effective approach. It promotes the use of multiple services and reduces the costs of each service (16). By integrating NCD care and RMNCH, both health services could be efficiently delivered; subsequently, this may help prevent the development of NCDs, especially during the perinatal period. In high-income countries, NCD care has been integrated into both the medical and public health services (e.g., cervical and breast cancer screening services) (17). This has effectively reduced the incidences of NCDs (18). In low- and middle-income countries, integrating health care services is also expected to be highly effective for NCDs. However, no prior systematic review article or protocol is available in this area, and the current integration status in these countries is unclear. Analyzing the effectiveness of integration can reveal

further opportunities for intervention and offer insights for tailoring interventions. Furthermore, it is necessary to closely examine the effectiveness and feasibility of NCD screening and treatment before adopting them as part of the health system, which can be expensive. In addition, integrating NCD care into regular RMNCH services might reduce the overall health care costs.

Therefore, in this review, we first aimed to quantitatively synthesize evidence regarding the effectiveness of integrating NCD interventions and RMNCH care services in low- and middle-income countries. Subsequently, we aimed to evaluate the impact of such integrations on the health outcomes and access to services for women and children.

2. Methods

We conducted a systematic review following the guidelines of the Cochrane Collaboration (19) and the four phases indicated in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Statement (20). We developed the systematic review based on the PRISMA checklist and registered the protocol for this systematic review at the PROSPERO international prospective register of systematic reviews on June 17, 2016 (registration number: CRD42015023425; available at <http://www.crd.york.ac.uk/PROSPERO>).

2.1. Study Inclusion Criteria

This review included randomized controlled trials and non-randomized intervention studies, which might provide evidence on the effects of interventions that cannot be randomized or adequately studied in randomized trials (19). We selected only peer-reviewed journal articles and reports from international organizations as publication types. We included studies in any language if they had English abstracts. Furthermore, we included studies only if they were conducted in low- and middle-income countries, as defined by the World Bank (21). We excluded non-intervention studies, such as case series, case reports, and qualitative studies, and those conducted in high-income countries from this review.

2.2. Participants

Participants of the interventions included women of reproductive age, defined as those aged 15-49 years according to the WHO; mothers; children < 5 years old; and RMNCH care service providers from low- and middle-income countries. We excluded studies that were conducted specifically in populations with an underlying disease or condition, such as in HIV-positive women, which would complicate the generalizability of our findings to a wider population.

2.3. Interventions and Controls

The included interventions comprised packaged care/services that integrated RMNCH services with NCD care. We focused on the following four NCD types: cardiovascular diseases, cancers, chronic respiratory diseases, and diabetes. We distinguished gestational and non-gestational NCDs, since gestational NCDs are often temporal and caused by pregnancy-induced changes in the hormonal levels. Anemia, domestic violence, depression, and diseases related to oral health are also often considered as NCDs (22-24), however, as the most frequently used WHO definition of NCDs includes cardiovascular diseases, cancers, chronic respiratory diseases, and diabetes (22), we only focused on these four types in this study. We considered studies in which RMNCH and NCD care were not integrated as the controls for this review.

2.4. Outcomes

We included studies with outcomes of maternal and/or infant mortality and complications, as well as health care service coverage, in this review.

2.5. Search Strategy

We used the following bibliographic databases to search for relevant articles: PubMed/Medline, POPLINE, EBSCO/CINAHL, BiblioMap, and ISI Web of Science. We reviewed relevant internet sources from the WHO library database and Google Scholar for additional gray literature. Finally, we identified additional studies using the snowball method by reviewing the reference lists of the retrieved articles. The publication period was set to 2000-2015 to ensure that enough studies were retrieved.

We conducted the literature search by using appropriate key words, accepted Medical Subject Headings words, and combinations thereof. One search approach involved using broad search terms (*e.g.*, mothers OR women OR female OR adolescent OR pregnant women) combined with intervention-specific search terms such as disease names (cardiovascular diseases OR cancer OR respiration disorders OR diabetes) and NCD risk factors (tobacco use OR smoking OR diet OR nutrition OR alcohol). To supplement the search, we also reviewed reports published by agencies, including the WHO, UNICEF, and the World Bank. Again, we conducted a further snowballing search through hand-searching or by searching for references from the initially identified studies.

2.6. Data Collection and Analysis

2.6.1. Selection of studies

The study selection process is summarized in Figure 1.

The first and second authors independently extracted data and screened the quality and content of the identified studies. We evaluated the titles and abstracts of the study articles for relevance and compliance with the selection criteria based on the research setting, study design, and reported outcomes. We then classified the articles as included, excluded, uncertain, or duplicate; all potentially included or uncertain studies were confirmed, and any disagreements were resolved by consensus.

2.6.2. Data extraction and management

The first author extracted the features of each study (*e.g.*, study design, setting, participants, components of intervention, related NCDs, and main outcomes) and entered them into a standardized form. The second author checked all data for accuracy and completeness. To resolve any discrepancies, the two authors discussed until they reached an agreement.

2.6.3. Risk of bias in the included studies

The first and second authors assessed the quality of the trials. For randomized controlled trials, the risk of bias was evaluated using the "risk of bias" tool presented in the Cochrane Handbook for Systematic Reviews of Interventions (19). Specifically, we assessed the risk of bias in the following seven domains: sequence generation, allocation concealment, blinding of participants and outcome assessors, incomplete outcome data, selective outcome reporting, and other potential threats to validity. To assess non-randomized intervention studies, we used "a Cochrane risk of bias assessment tool: for non-randomized studies of interventions (ACROBAT-NRSI)" (25). For non-randomized intervention studies, we evaluated the risk of bias using the following seven domains: confounding, selection of participants, intervention measurements, departures from intended interventions, missing data, outcome measurement, and selective outcome reporting. In addition, for non-randomized intervention studies, we assessed the overall risk of bias following the results of the seven domains. We resolved all discrepancies by consensus.

2.6.4. Analysis

We analyzed the results only narratively because the outcome measurements and interventions differed among the selected studies. First, we summarized the included study interventions according to the related NCDs, settings, study design, participants, intervention components, outcomes, and quantitative results. Second, we stratified the studies according to the NCD type integrated with RMNCH care and provided narrative descriptions for each NCD type.

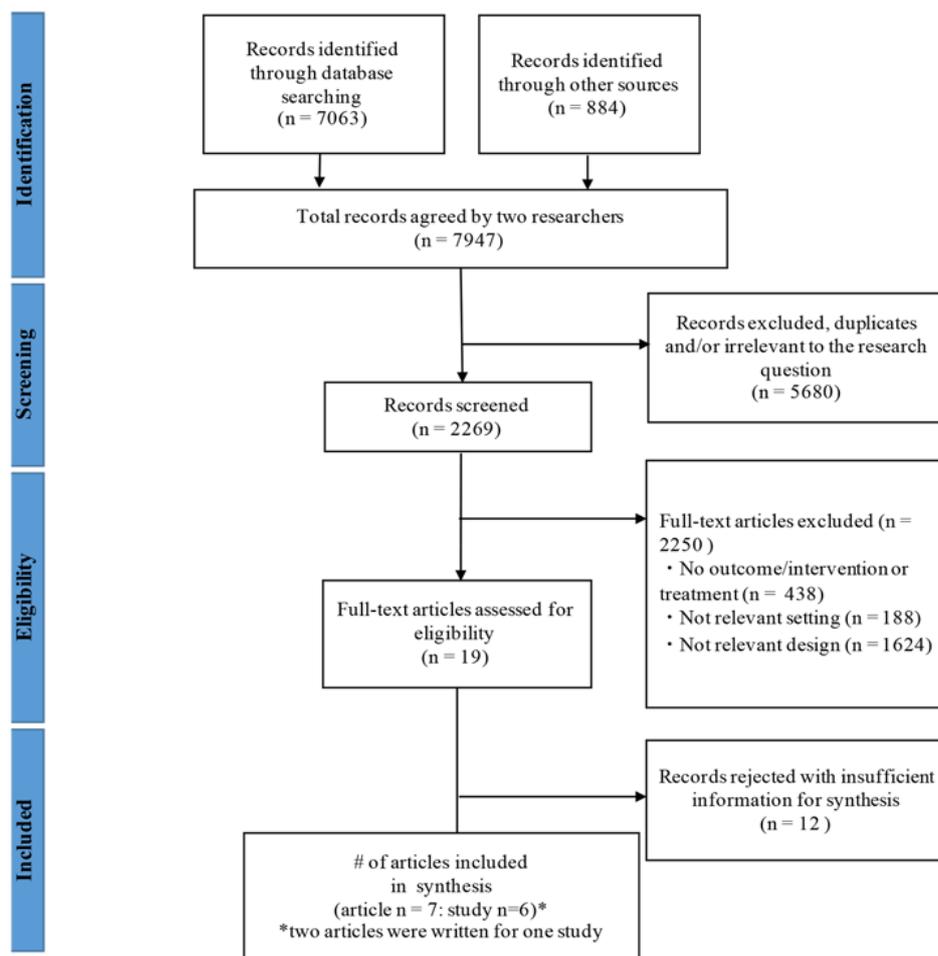


Figure 1. Diagram of information flow through phases of systematic review.

3. Results

A total of 7,947 articles were obtained from the following sources: PubMed ($n = 7,063$), CINAHL ($n = 353$), CENTRAL ($n = 295$), Web of Science ($n = 167$), PsycINFO ($n = 62$), POPLINE ($n = 4$), and Scopus ($n = 3$). After the initial screening, we included 2,269 studies, because the other studies were duplicates and/or irrelevant to the research questions. Of these studies, we excluded a further 2,250 because they had no outcome/intervention or treatment ($n = 438$), irrelevant setting ($n = 188$), or irrelevant design ($n = 1624$). Subsequently, we evaluated the remaining 19 studies for eligibility by reviewing the full-text articles. Among them, we included two randomized controlled studies (three published articles) and four non-randomized interventions (26-32), which were all written in English and had been published in peer-reviewed journals, in the final analysis. For randomized controlled trial studies, we initially selected three articles. However, we integrated those written by Jayashree *et al.* and Rakhshani *et al.* into the same category (*i.e.*, the Jayashree-Rakhshani study), because both articles involved the same intervention

study (27, 32).

3.1. Risk of Bias Assessment

Table 1 shows the risks of bias for the two randomized controlled trial studies. Both assessed studies lacked information related to the blinding during data collection or implementation. All other biases were "low" in the selected studies. However, high attrition bias was observed in the Jayashree-Rakhshani study, because the dropout rate was almost 35% of the randomized participants in the intervention group (27,32).

Table 2 shows the risk of bias in the four non-randomized intervention studies. The overall risks of bias were "moderate" in the studies by Moon *et al.* and Aalami *et al.* (26,30), and "serious" for the studies by Plotkin *et al.* and Luo *et al.* (29,31). In Plotkin's study, we assessed the "bias due to missing data" as "serious" because 40% of the participants were not provided all interventions. As for Luo's study, we assessed the "bias in selection of participants" as "serious" because the allocation of the intervention and control groups was not indicated.

Table 1. Risk of bias in the included randomized controlled trials

Risk of bias	Jayashree 2013, Rakhshani 2012	Kasawara 2013
Random sequence generation	Low	Low
Allocation concealment	Low	Low
Blinding of participants and personnel	Low	Low
Blinding of outcome assessment	Unclear	Unclear
Incomplete outcome data	High	Low
Selective reporting	Low	Low
Other bias	Low	Low

Table 2. Risk of bias in the included non-randomized intervention studies

Risk of bias	Aalami 2016	Luo 2014	Moon 2012	Plotkin 2014
Bias due to confounding	Moderate	Moderate	Moderate	Low
Bias in selection of participants into the study	Moderate	Serious	Moderate	Low
Bias in measurement of interventions	Moderate	No information	Moderate	Low
Bias due to departures from intended interventions	Low	Low	Low	Low
Bias due to missing data	Low	Low	Low	Serious
Bias in measurement of outcomes	Moderate	Moderate	Low	Low
Bias in selection of the reported results	Moderate	Moderate	Moderate	Moderate
Overall risk of bias	Moderate	Serious	Moderate	Serious

3.2. *Randomized Controlled Trials Integrating RMNCH Care and NCD Interventions*

We included two randomized controlled studies, conducted by Jayashree-Rakhshani in India and Kasawara in Brazil (Table 3) (27,28). The study interventions focused on gestational hypertension and diabetes. We did not identify any interventions with a focus on NCDs in general, women-specific cancers, or pediatric NCDs. The total number of participants was 184, including both the intervention and control groups. Both studies involved high-risk pregnant women and physical exercise interventions (*i.e.*, yoga intervention in Jayashree-Rakhshani's study and bicycle exercises in Kasawara's study).

3.3. *Non-randomized Controlled Trials Integrating RMNCH and NCD Care*

We identified four non-randomized controlled trials in this review (Table 3). These studies were conducted by Aalami et al. in Iran, Luo et al. in China, Moon et al. in Mozambique, and Plotkin et al. in Tanzania (26,29-31). The study interventions focused on gestational diabetes and cervical cancer. We were unable to identify studies that focused on NCDs in general or pediatric NCDs. All the selected studies were intervention studies; however, control groups were only identified in the studies by Luo et al. and Aalami et al (26,29). In the remaining two studies, the researchers evaluated the implementation and intervention of a health care program, without a separate control group (30,31).

3.4. *Integration of RMNCH Care according to the NCD Type*

As indicated in Table 4, we divided the identified studies into three categories according to the type of NCD care that was integrated with the RMNCH care. Of note, the studies on cardiovascular disease and diabetes focused only on gestational hypertension or diabetes, and not on cardiovascular disease or diabetes as a general chronic disease. Cervical cancer was the only cancer that was linked with RMNCH care in the selected studies. Most of the health care interventions for NCDs occurred during the mothers' pregnancy period, except for cervical cancer, which was integrated into the RMNCH service (*i.e.*, during family planning and HIV tests) before or after the pregnancy period.

3.5. *Effectiveness of Integrating NCD Interventions and RMNCH Care on Maternal Health*

3.5.1. *Gestational NCDs*

Four studies evaluated the integration between gestational hypertension and RMNCH care (26-29). In the studies by Jayashree-Rakhshani and Kasawara, a physical exercise intervention was implemented for pregnant women. In the studies by Aalami et al. and Luo et al., the intervention involved breathing techniques and nutritional management, respectively. The effectiveness of these interventions on maternal health differed. The progressive muscular relaxation ($p < 0.01$) and breathing interventions ($p < 0.01$) significantly reduced blood pressure in one study (26). The incidence of preeclampsia was significantly reduced by exercise intervention in the Jayashree-Rakhshani study (absolute risk reduction, 21%) and by nutrition management intervention ($p < 0.05$) (27,29,32). However, we did not observe significant

Table 3. Characteristics of the included interventions that integrated non-communicable disease interventions and maternal, newborn, and child health care services

Related non-communicable diseases	Author/country	Study design	Participants	Intervention	Outcomes
Gestational hypertension (cardiovascular disease)/ gestational diabetes	Jayashree <i>et al.</i> 2013 and Rakshani <i>et al.</i> 2012/India (27,32)	RCT	Women with high-risk pregnancy at 12 weeks' gestation ($n = 68$; 30 interventions and 38 controls). The inclusion criteria were as follows: high-risk obstetric history, twin pregnancies, maternal age < 20 or > 35 years, obesity (BMI > 30), and genetic history of pregnancy complications.	The yoga group received standard care plus 1-hour yoga sessions 3 times a week from the beginning of the 13th week to the end of the 28th week of gestation. The control group received standard care plus 30-minute walks every morning and evening within the same period. The yoga classes were conducted by well-trained certified yoga therapists. Three categories of yoga interventions were offered to the study group: 1) breathing exercises, 2) yoga postures, and 3) meditative exercises. The latter included visualization, guided imagery, and sound resonance techniques. Subjects were asked to visualize the fetus in the uterus, umbilical cord, and placenta.	Platelet count, uric acid level, gestational hypertension, preeclampsia, and intrauterine growth restriction
Gestational hypertension (cardiovascular disease)	Kasawara <i>et al.</i> 2013/Brazil (28)	RCT	Pregnant women with chronic hypertension and/or previous preeclampsia, considered at high risk of preeclampsia development ($n = 116$; 58 interventions and 58 controls)	Participants were divided into two groups: (1) a study group that performed physical exercise with a stationary bicycle once a week for 30 minutes with controlled intensity (i.e., heart rate 20% above resting values) under professional supervision and (2) a control group that was not engaged in any physical exercise. Women in the study group performed 9.24 (± 7.03) physical exercise sessions.	Maternal outcomes: admission to the ICU and morbid conditions due to complications. Infant outcomes: birth weight, Apgar score, admission to ICU, and infant morbidity.
Gestational hypertension (cardiovascular disease)	Adami <i>et al.</i> 2016/Iran (26)	Non-randomized intervention study	Pregnant women after 20 weeks of gestational age with systolic BP ≥ 135 mmHg or diastolic BP ≥ 85 mmHg ($n = 60$; assigned to a progressive muscular relaxation group, breathing control exercises group, or control group)	Progressive muscular relaxation and breathing control exercises were administered to two experimental groups in person once a week and by instructions given on a CD for the remaining days for 4 weeks. Participants were followed-up regarding continuation of exercises by phone calls. The symptoms of severe preeclampsia were evaluated by phone calls and at the beginning of each personal class. In the personal classes, necessary explanations about performing exercises after the primary interventions and controlling blood pressure were provided. Blood pressure was evaluated before and after the interventions in the intervention group and before and after 15 minutes of waiting without any special intervention in the control group.	Systolic blood pressure, diastolic blood pressure
Gestational hypertension (cardiovascular disease)/ gestational diabetes	Luo <i>et al.</i> 2014/China (29)	Non-randomized intervention study	Pregnant women ($n = 276$; 131 in the intervention group and 145 in the control group)	Women in the intervention group received individualized nutritional management in addition to routine prenatal care. They received monthly obstetric visits up to week 28, biweekly visits from week 28 to week 36, and weekly visits until the end of pregnancy. The total energy intake was modified according to the progression of pregnancy every 4 weeks with energy requirements being individually estimated according to each mother's BMI, weight, and obstetric condition. Women in the control and intervention groups received the same initial information regarding the purpose and content of this study.	Gestational diabetes, gestational weight gain, birth weight, Apgar score, and incidence of pregnancy complications
Cervical cancer	Moon <i>et al.</i> 2012/Mozambique (30)	Evaluation study of health care program implementation	Screened women ($n = 4651$)	Cervical cancer screening was nested into family planning clinics using VIA by maternal child health nurses. During the examination, women were evaluated for clinical signs of sexually transmitted infections and other gynecologic pathologies. A cotton swab was then applied, and excess mucus was removed prior to applying 5% acetic acid. After 1 minute, the findings were reported as positive or negative for possible squamous intraepithelial lesions. Women with lesions involving $> 75\%$ of the transformation zone, endocervical involvement, or suspected invasive cancer were referred to hospital for loop electrosurgical excision procedure, biopsy, or surgery.	Cervical cancer screening
Cervical cancer	Plotkin <i>et al.</i> 2014/Tanzania (31)	Evaluation study of health care program implementation, analysis of the 2010–2013 data	Screened women ($n = 24,966$) collected from routine client level service delivery data from 21 government health facilities	Integrated HIV testing into cervical cancer screening with VIA. Providers working at the onsite HIV CTCs at facilities received a 1-day orientation on cervical cancer screening, emphasizing the importance of referral for HIV-infected women. Cervical cancer screening and HIV testing were provided during the same visit and in the same location. Cervical cancer screening was offered at the RCH clinic at least twice a week. Family planning services were offered through the RCH and were available every day when cervical cancer screening was offered. If a client was reactive to an HIV test administered during the screening, she was referred to the onsite HIV CTC.	Acceptance of the HIV test by clients and the number and proportion of women being offered the HIV test

BMI, body mass index; ICU, intensive care unit; RCT, randomized controlled trial; CD, compact disc; BP, blood pressure; VIA, visual inspection with acetic acid; HIV, human immunodeficiency virus; CTC, care and treatment center; RCH, reproductive and child health.

Table 4. Main findings regarding maternal and child health outcomes stratified by the non-communicable disease types

Related non-communicable diseases	Related study	Main findings for mothers	Main findings for children
Gestational hypertension (cardiovascular disease)	Jayashree-Rakhshani/India (27)	<ul style="list-style-type: none"> Significantly fewer women in the yoga group ($n = 3$) developed gestational hypertension/preeclampsia than those in the control group ($n = 12$; absolute risk reduction, 21%). No significant difference in platelet count in the intervention group No significant difference in the increase in uric acid level in both groups No significant differences of eclampsia between the intervention and control groups No significant difference in preeclampsia and other maternal morbidity incidences between the groups 	<ul style="list-style-type: none"> Significantly fewer SGA newborns were born in the yoga group ($p = 0.03$). Significantly fewer babies with low 1-minute and 5-minute Apgar scores were born in the yoga group ($p = 0.01$ and $p = 0.04$). No significant difference in birth weight, Apgar scores, or other incidences of neonatal morbidity between the groups.
	Kasawara/Brazil (28)	<ul style="list-style-type: none"> After 4 weeks of intervention, significant differences in BP in the progressive muscular relaxation group ($p < 0.001$ for systolic BP and $p = 0.001$ for diastolic BP) and breathing in the control group ($p = 0.004$ for systolic BP and $p = 0.047$ for diastolic BP) 	<ul style="list-style-type: none"> No significant differences in birth weight, infant length, gestational age, and major congenital anomalies between the groups.
	Alami/Iran (26)	<ul style="list-style-type: none"> Significant difference in the preeclampsia incidence rate: 3.1% (intervention) vs. 11% (control) $p < 0.05$ No significant difference in the incidences of preterm labor ($p = 0.736$) and premature membranes rupture 	<ul style="list-style-type: none"> Significantly fewer cases of intra-uterine growth restriction in the intervention group (6.9%) than the control group (25.8%) ($p = 0.05$). No significant differences in birth weight, infant length, and gestational age between the groups. Significant difference in macrosomia ($p = 0.026$).
Gestational diabetes	Luo/China (29)	<ul style="list-style-type: none"> Significantly fewer gestational diabetes cases in the intervention group (3.4%) than in the control group (20.0%) ($p = 0.05$) 	<ul style="list-style-type: none"> No significant differences in birth weight, infant length, and gestational age between the groups.
	Jayashree-Rakhshani/India (27)	<ul style="list-style-type: none"> Significant difference in gestational diabetes: 3.8% (intervention) vs. 14.5% (control) ($p = 0.002$) 	<ul style="list-style-type: none"> No significant differences in birth weight, infant length, and gestational age between the groups.
Cervical cancer	Luo/China (29)	<ul style="list-style-type: none"> Significant difference in gestational weight gain: 7.58 kg (intervention) vs. 12.57 kg (control) ($p < 0.01$) VIA was judged positive for squamous intraepithelial lesions in 8% of women ($n = 380$). About 4% of these women had lesions requiring referral to hospital; 14 of them were examined at the hospital, but their records were inadequate to judge the outcomes. Among the screened women, 2714 either had knowledge of their HIV status prior to the VIA or were subsequently sent for HIV testing; 583 of these women were found to be HIV positive. About 96% of women eligible for cryotherapy received treatment on the day of the VIA screen, as compared with only 53% in the first quarter of the year. 	<ul style="list-style-type: none"> Approximately 26% were referred from HIV care and treatment clinics. Among women of unknown HIV status ($n = 18,539$), 60% were offered an HIV test. Of those offered testing, 94% of women accepted the testing.
	Moon/Mozambique (30)	<ul style="list-style-type: none"> Approximately 26% were referred from HIV care and treatment clinics. Among women of unknown HIV status ($n = 18,539$), 60% were offered an HIV test. Of those offered testing, 94% of women accepted the testing. 	<ul style="list-style-type: none"> Approximately 26% were referred from HIV care and treatment clinics. Among women of unknown HIV status ($n = 18,539$), 60% were offered an HIV test. Of those offered testing, 94% of women accepted the testing.
	Plotkin/Tanzania (31)	<ul style="list-style-type: none"> Approximately 26% were referred from HIV care and treatment clinics. Among women of unknown HIV status ($n = 18,539$), 60% were offered an HIV test. Of those offered testing, 94% of women accepted the testing. 	<ul style="list-style-type: none"> Approximately 26% were referred from HIV care and treatment clinics. Among women of unknown HIV status ($n = 18,539$), 60% were offered an HIV test. Of those offered testing, 94% of women accepted the testing.

SGA, small for gestational age; BP, blood pressure; VIA, visual inspection with acetic acid; HIV, human immunodeficiency virus.

differences in the blood pressure levels and incidence of preeclampsia in the physical intervention study by Kasawara et al (28). Further, we observed significant differences in the incidence of eclampsia, platelet count, and uric acid levels in the Jayashree-Rakhshani study (27,32).

Studies by Jayashree-Rakhshani's and Luo's evaluated the integration of gestational diabetes and RMNCH care. According to these studies, exercise intervention significantly reduced the incidence of gestational diabetes ($p = 0.05$) (27,32). Moreover, nutritional management intervention significantly reduced the incidences of gestational diabetes ($p < 0.01$) and gestational weight gain ($p < 0.01$) (29).

3.5.2. Cancer

Two studies evaluated the integration of cervical cancer and RMNCH care. Those studies aimed to examine the feasibility of cervical cancer screening by visual inspection with acetic acid (30,31). However, no clear maternal health outcome was reported in those studies, because the researchers were more focused on the uptake of the screening service. According to Moon et al., the incidences of positive cervical cancer cases (8.0%) were comparable to those of similar interventions (range, 3.8-12.7%) (30). Among the women screened for cervical cancer in Moon et al. and Plotkin et al.'s studies, those with unknown HIV status were also tested for HIV (58% and 47% of screened women, respectively), and 21% and 26% were identified as HIV positive, respectively (30,31).

3.6. Effectiveness of Integrating NCD Interventions and RMNCH Care on Child Health

In the yoga intervention group in the study by Jayashree-Rakhshani, significantly fewer babies were born with low 1-minute and 5-minute Apgar scores ($p = 0.01$ and $p = 0.04$, respectively), and significantly fewer fetuses had intra-uterine growth restriction (27,32). In Kasawara *et al.*'s bicycle exercise intervention, no significant difference was observed in the fetus/newborn outcomes (*i.e.*, birth weight and Apgar scores) compared with those in the control group (28). Moreover, in the study by Luo *et al.*, no significant difference was observed in the birth weight and infant length in the nutritional management intervention (29).

4. Discussion

This is the first systematic review to examine the effectiveness of integrating gestational and non-gestational non-communicable disease interventions and antenatal care among women and children in low- and middle-income countries. We demonstrated some new and important findings. First, the included studies

focused only on gestational hypertension and diabetes, and on cervical cancer. Second, among the interventions integrating NCD and RMNCH care, yoga exercise and nutritional improvements were effective in reducing gestational hypertension and diabetes. Third, studies that linked cervical cancer and RMNCH care revealed positive impacts on the uptake of cervical cancer screening services and on identifying potential cervical cancer cases. However, in this systematic review, we included only two randomized controlled trials, and we were unable to conduct a meta-analysis for these studies because the nature of the interventions differed between the studies (*i.e.*, yoga *vs.* bicycle exercise) and because of the small sample sizes. Thus, more concrete data are needed to fully understand the effectiveness of integrating NCD interventions and RMNCH care in low- and middle-income countries.

According to this review, yoga and nutritional improvements were effective in reducing gestational hypertension as well as gestational diabetes. These interventions would not incur high costs and, therefore, would be feasible in low- and middle-income countries. However, the results of this review are not fully generalizable, owing to the insufficient quality and sample size. Thus, further intervention studies are needed to confirm our findings.

Furthermore, the integration of cervical cancer screening and RMNCH care showed a positive impact on the uptake of cervical cancer screening services and on identifying potential cervical cancer cases. In general, integrating cervical cancer screening into the regular RMNCH services is considered costly in low- and middle-income countries. However, using visual inspection with acetic acid would be affordable in these countries and would help improve the health indicators (33). With further evidence, the integration of cervical cancer screening and RMNCH care could be accelerated.

In this review, we did not identify any studies that linked RMNCH care and breast cancer, which has the highest mortality rate among all cancers in women in resource-limited settings (*i.e.*, low- and middle-income countries) (34). Breast cancer screening using mammography/ultrasound is considered effective, but is not realistic in resource-limited settings. Therefore, it is important to examine the affordability of other possible screening methods (*e.g.*, palpation or methods that target only high-risk groups) and the effectiveness of integrating these methods with RMNCH programs.

In addition, we did not identify any studies that integrated pediatric NCD and RMNCH care. Respiratory disease (*e.g.*, pediatric asthma) is a key cause of mortality in children of low-income countries (35). Therefore, the child mortality rate would theoretically improve by integrating respiratory disease care into the regular health care for pediatric patients. Studies on preventing passive smoking among women

have been conducted in a resource-limited setting, but the intervention was not linked with RMNCH care (36,37). Thus, further research is required to examine the effectiveness of integrating respiratory disease and RMNCH care on child health indicators.

Recommendations for Integrating NCD and RMNCH Services To plan a strategy for integrating NCD and RMNCH care services, gestational NCD and other NCD care services should be considered separately. Gestational NCD care services could be integrated into pre-pregnancy or antenatal care through the exercise or nutritional interventions that were suggested as effective in this review. Other NCD care services, including screening for cervical or breast cancer, could be integrated into the postnatal care or family planning services. By integrating these services, health workers can provide different services at the same time and at a low cost, thereby avoiding overlooking many common health problems of women. However, such integration may result in the focus of the service being diluted (38). Therefore, health workers need to be allocated and trained adequately, and continuous monitoring is required to assess whether the care can be provided to all women with the same effect (16). Moreover, to maximize the effects of integration, the potential challenges need to be examined in detail.

Limitation The main limitation of this study was that the lack of randomized studies, which might have obscured the actual results of the systematic review. Further evidence is needed to determine the effectiveness of integrating NCD interventions and RMNCH care.

5. Conclusions

Interventions that integrate NCDs and RMNCH care appear effective in reducing gestational hypertension and diabetes in low- and middle-income countries. Furthermore, integrating cervical cancer screening and RMNCH care showed a positive impact for identifying suspected cervical cancer cases. However, primary evidence remains scarce, and further research is required to conclude the effectiveness of integrating NCD interventions and RMNCH care services.

Funding This research was supported by Kyushu University.

References

1. WHO. Health in 2015: from MDGs, millennium development goals to SDGs, sustainable development goals. Geneva, Switzerland, WHO. 2015.
2. Alliance N. Non-communicable diseases: A priority for women's health and development. Brussels, Belgium, NCD Alliance. 2011.
3. WHO. Integration of the prevention and control of non-communicable diseases and maternal and newborn health programmes and services – a case for action. Vol. 2017.

4. WHO. Global health risks: Mortality and burden of disease attributable to selected major risks. Geneva, Switzerland, WHO. 2009.
5. The Partnership for Maternal Newborn and Child Health. PMNCH Knowledge Summaries: #15 - Non-communicable diseases. Geneva, Switzerland, The Partnership for Maternal, Newborn and Child Health. 2011.
6. Porter P. "Westernizing" women's risks? Breast cancer in lower-income countries. *N Engl J Med.* 2008; 358:213-216.
7. Gyawali B, Shimokata T, Honda K, Tsukuura H, Ando Y. Should low-income countries invest in breast cancer screening? *Cancer Causes Control.* 2016; 27:1341-1345.
8. Varughese J, Richman S. Cancer care inequity for women in resource-poor countries. *Rev Obstet Gynecol.* 2010; 3:122-132.
9. NCD Alliance. Children in every policy: Recommendation for a lifecourse approach to NCDs. NCD Alliance. 2011.
10. UN. The Millennium Development Goals Report 2010. New York, USA, UN. 2010.
11. Pattinson R, Kerber K, Buchmann E, Friberg IK, Belizan M, Lansky S, Weissman E, Mathai M, Rudan I, Walker N, Lawn JE; Lancet's Stillbirths Series steering committee. Stillbirths: How can health systems deliver for mothers and babies? *Lancet.* 2011; 377:1610-1623.
12. National Collaborating Centre for Women's and Children's Health. Antenatal care: Routine care for the healthy pregnant woman. London, UK. National collaborating centre for women's and children's health. 2008.
13. Mosca L, Benjamin EJ, Berra K, *et al.* Effectiveness-based guidelines for the prevention of cardiovascular disease in women – 2011 update: A guideline from the American Heart Association. *J Am Coll Cardiol.* 2011; 57:1404-1423.
14. Eriksson J, Lindstrom J, Tuomilehto J. Potential for the prevention of type 2 diabetes. *Br Med Bull.* 2001; 60:183-199.
15. UN. Global strategy for women's and children's health. New York, USA, UN, Every Woman Every Child. 2010.
16. Maina WK. Integrating noncommunicable disease prevention into maternal and child health programs: Can it be done and what will it take? *Int J Gynaecol Obstet.* 2011; 115(Suppl 1):S34-36.
17. Sahasrabudde VV, Parham GP, Mwanahamuntu MH, Vermund SH. Cervical cancer prevention in low- and middle-income countries: Feasible, affordable, essential. *Cancer Prev Res (Phila).* 2012; 5:11-17.
18. Forouzanfar MH, Foreman KJ, Delossantos AM, Lozano R, Lopez AD, Murray CJ, Naghavi M. Breast and cervical cancer in 187 countries between 1980 and 2010: A systematic analysis. *Lancet.* 2011; 378:1461-1484.
19. Higgins J, Green S. *Cochrane handbook for systematic reviews of interventions.* NJ, USA. Wiley, 2008.
20. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *BMJ.* 2009; 339:b2535.
21. World Bank country and lending groups. <http://data.worldbank.org/about/country-and-lending-groups>. (accessed April 15, 2018)
22. Noncommunicable diseases. <http://www.who.int/mediacentre/factsheets/fs355/en/> (accessed April 15,

- 2018)
23. Noncommunicable diseases. <http://www.afro.who.int/en/noncommunicable-diseases/> (accessed April 15, 2018)
 24. Noncommunicable diseases. <http://www.euro.who.int/en/health-topics/noncommunicable-diseases> (accessed April 15, 2018)
 25. Sterne JA, Hernan MA, Reeves BC, *et al.* ROBINS-I: A tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*. 2016; 355:i4919.
 26. Aalami M, Jafarnejad F, ModarresGharavi M. The effects of progressive muscular relaxation and breathing control technique on blood pressure during pregnancy. *Iran J Nurs Midwifery Res*. 2016; 21:331-336.
 27. Jayashree R, Malini A, Rakhshani A, Nagendra H, Gunasheela S, Nagarathna R. Effect of the integrated approach of yoga therapy on platelet count and uric acid in pregnancy: A multicenter stratified randomized single-blind study. *Int J Yoga*. 2013; 6:39-46.
 28. Kasawara KT, Burgos CS, do Nascimento SL, Ferreira NO, Surita FG, Pinto E, Silva JL. Maternal and perinatal outcomes of exercise in pregnant women with chronic hypertension and/or previous preeclampsia: A randomized controlled trial. *ISRN Obstet Gynecol*. 2013; 2013:857047.
 29. Luo XD, Dong X, Zhou J. Effects of nutritional management intervention on gestational weight gain and perinatal outcome. *Saudi Med J*. 2014; 35:1267-1270.
 30. Moon TD, Silva-Matos C, Cardoso A, Baptista AJ, Sidat M, Vermund SH. Implementation of cervical cancer screening using visual inspection with acetic acid in rural Mozambique: Successes and challenges using HIV care and treatment programme investments in Zambezia Province. *J Int AIDS Soc*. 2012; 15:17406.
 31. Plotkin M, Besana GV, Yuma S, Kim YM, Kulindwa Y, Kabole F, Lu E, Giattas MR. Integrating HIV testing into cervical cancer screening in Tanzania: An analysis of routine service delivery statistics. *BMC Womens Health*. 2014; 14:120.
 32. Rakhshani A, Nagarathna R, Mhaskar R, Mhaskar A, Thomas A, Gunasheela S. The effects of yoga in prevention of pregnancy complications in high-risk pregnancies: A randomized controlled trial. *Prev Med*. 2012; 55:333-340.
 33. Denny L, de Sanjose S, Mutebi M, Anderson BO, Kim J, Jeronimo J, Herrero R, Yeates K, Ginsburg O, Sankaranarayanan R. Interventions to close the divide for women with breast and cervical cancer between low-income and middle-income countries and high-income countries. *Lancet*. 2017; 389:861-870.
 34. Stewart B, Wild C. World cancer report 2014. Geneva, Switzerland, IARC. 2014.
 35. WHO. The world health report 2002 – reducing risks, promoting healthy life. Geneva, Switzerland, WHO. 2002.
 36. Nichter M, Nichter M, Padmawati RS, Ng N. Developing a smoke free household initiative: An Indonesian case study. *Acta Obstet Gynecol Scand*. 2010; 89:578-581.
 37. Nichter M, Padmajam S, Nichter M, Sairu P, Aswathy S, Mini GK, Bindu VC, Pradeepkumar AS, Thankappan KR. Developing a smoke free homes initiative in Kerala, India. *BMC Public Health*. 2015; 15:480.
 38. Dehne KL, Snow R, O'Reilly KR. Integration of prevention and care of sexually transmitted infections with family planning services: What is the evidence for public health benefits? *Bull World Health Organ*. 2000; 78:628-639.

(Received April 17, 2018; Revised April 23, 2018; Accepted April 24, 2018)

The impact of parental migration on injuries among left behind young people aged 10 years to 24 years in Botswana

Lesego Selotlegeng^{1,2,*},

¹ United Nation Economic Commission for Africa, Addis Ababa, Ethiopia;

² Institute of Development Management Idm (Botswana), Gaborone, Botswana.

Summary

There has been little evidence of the relationship between children and absence of parents in Botswana literature; and it is still the case that absence of parents increases the risk of injuries to their children. The aim of this study is to investigate the impact of parental migration among left behind young people aged 10 years to 24 years in Botswana and examine patterns of injuries from immigrant families. This is a population-based cross-section from which a dataset was created by compiling data from two different sources for the period of 2010 to 2015. All the variables were transformed into natural logarithms so as to avoid outliers and normalize the variables. All statistical modeling was carried out using Statistical Analysis System version 6.1. Left behind young people were associated with higher risk of injuries especially in motor vehicle accidents 391(40%), fire 264 (27%), and burns 162 (16%) except for drowning, machinery and poisoning. Parental migration is the most important issue in the total number of injuries. Correlation coefficient shows that non-left behind, the greater the chance to be classified under the poisoning cases ($r = 0.888$) and is lower compared to the left behind ($r = 0.471$). This study shows clearly that injuries take an insufficiently high toll on children's health and on society. Additionally, if parental migration trends continue, Botswana's burden of injuries are expected to rise in the next 10 years.

Keywords: Parental migration, young people, injuries, Botswana

1. Introduction

Migration is a complex phenomenon, often viewed in the context of benefits or risks. Despite the growing public discourse on the subject of migration, little is known about the migration process itself, notably, what the feelings of their family members left behind are and the effects of their migration. In African countries, parents are primarily responsible for the care of their children. However, out-migration of parents is common in these countries. Migration can have an impact on mental, emotional health, physical, and wellbeing of those left behind especially young people in the place of origin. The connection between migration and health status

is duplex, in which the act of migration may influence health outcomes of those left behind (1). Parental Migration might also unfavorably affect the young people's levels of satisfaction leading to loneliness that might aggravate emotional wellbeing (2).

The generation between 10 and 24 years holds a lot of potential yet faces a lot challenges, especially when their parents migrated in search of better opportunities. There is little consistency in the young people definitions across countries and organizations when it comes to using an age-based definition. For analytical purposes, the World Health Organization (WHO) defines young people as individuals between ages 10 and 24 (3). Migration is emerging as one of the key issues affecting left behind young people. Since their parents are increasingly migrating in search of survival, security, improved standards of living and education. The United Nation's report on migrants between 24 and 49 years of age, concerns who in 2013 represented about one-eighth (28.2 million) of the 232 million international migrants worldwide (4).

*Address correspondence to:

Dr. Lesego Selotlegeng, United Nation Economic Commission for Africa, Menelik II Ave. P.O. Box 3001, Addis Ababa, Ethiopia. Or Institute of Development Management Idm (Botswana). PO Box 1357, Gaborone, Botswana.
E-mail: tshepisokgetse@gmail.com

Injuries are defined as damage to a person caused by an acute transfer of mechanical, electrical, chemical or radiation energy (5). Worldwide, injuries are the second leading cause of death among young people accounting for 13% of the total years lost because of disabilities in this cohort (6). Nearly every day 2,300 adolescents die from injuries sustained from motor vehicle injuries, drowning, poisoning, falls, burns, and violence, while motor vehicle injuries alone are responsible for 10.2 deaths per 100,000 adolescents. More than 3,400 individuals die on the world's roads every day. In addition, tens of millions of people are injured or disabled every year. Children, pedestrians and older people are among the most vulnerable of road users. Drowning is a leading killer, almost 360,000 people lost their lives to drowning in 2015, and drowning is the third leading cause of death worldwide for children aged 5-14 years. Over 90% of drowning deaths occur in low- and middle-income countries. Burns and falls are significant causes of morbidity and mortality all over the world. They are among the leading causes of death among children under the age of 15 years (7). Parental migration has a negative effect on children's emotions, which may contribute to the cause of injuries among these children.

Botswana has been undergoing rapid urbanization in the last few decades. By 2011, over 100 thousand people in Botswana were migrants (8). The majority of internal migrants are adults aged 24-49 years old, moving from rural to urban areas to seek greener pastures (9). Children often do not move with their parents due to financial constraints and nature of work in urban areas. The term left behind young people referred to people (aged between 10 years to 24 years) in their rural hometowns and villages whose parents migrated to urban areas or internationally. By 2011, the number of young people living in rural areas across Botswana was estimated to be more than fifty thousand (10). Injuries are one of the leading causes of death and hospitalization for children and youth in Botswana (11). Those left behind children, youth in Canada have a higher risk of injuries than the children and youth whose parents did not migrate (12). The previous study in China indicated that left behind adolescents encounter educational problems, such as higher rates of drop outs (13). In this study, we investigated the impact of parental migration among left behind young people in Botswana and examined patterns of injuries in young people from immigrant families. We were able to describe the impact of those left behind, through comparison with their counterparts in the same rural area whose parents had not migrated.

2. Materials and Methods

2.1. Materials

This is population-based cross-section from which a dataset was created by compiling data from two

different sources for the period of 2010 to 2015. Ministry of Health in Botswana provides standard data collection to capture hospitalization and emergency department visits and includes the main diagnoses for the period of 2010 to 2015. These were used to obtain injury-related information where emergency department visits and hospitalization occurred. The department of civil and national registration provided the death dataset, which was used to identify in and out of hospital injury-related deaths in Botswana. Botswana has a population of 2,319,295. The Gross Domestic Product (GDP) in Botswana was worth 15.27 billion US dollars in 2016. The GDP value of Botswana represents 0.02 percent of the world economy. GDP in Botswana averaged 4.46 USD Billion from 1960 until 2016, reaching an all time high of 16.26 USD Billion in 2014 and a record low of 0.03 USD Billion in 1961 (14). Botswana's extraordinarily high inequality is holding the country back, making it difficult for sustained growth to lead to rapid poverty reduction. Poverty remains high in rural areas, female-headed households, and those with low levels of education. Alarming, poverty is concentrated among children and youth, with significant implications for social inclusion and inter-generational effects (15).

2.2. Study population

In order to be consistent with the national definition, this study adopted WHO definition of young people "Children between 10 years to 24 years who are left behind at their original residence, while one or both parents migrate into other places and have not been living with them for at least six months. This study uses two categories when it comes to age. Young people refers to 10-14 years and "old people" which refers to 15-24 years. Separation due to divorce or deceased parents did not constitute left behind young people. Information about young people's gender, age, education, occupation of parents, who are they living with; injury-related visits to emergency department and hospitalization was collected. Sample size is 2991.

2.3. Statistical analysis

Descriptive statistics were performed for the main independent variables. The panel data has been used for bivariate and multivariate linear regression analysis. Two dependent variables have been used for multivariate linear regression analysis (for example *i*) "left behind young people"; *ii*) "non-left behind young people"). These responses were considered as polytomous variables, with these two categories as possible answers. Explanatory variables were selected based on theoretical considerations. To compare outcomes of immigrants by migration status, multiple variables Poisson regression models adjusting for over

dispersion were used to compute rate ratios with 95% confidence intervals. For each model, variables were selected a priori and included in the regression analysis. The following function was developed to estimate the effect of migration on injuries.

$$\text{Injuries} = \beta_0 + \beta_1 \text{FB}_t + \beta_2 \text{NFB}_t + \mu_t$$

Where, injuries refers to poisoning, falls, drowning, cuts/piercing, motor road traffic, burns, suffocation, machinery and fires. FB refers to those that were left behind while NFB is non-left behind young people aged 10 years to 24 years and the stochastic error term that is assumed to be distributed with a zero mean. All variables were transformed into natural logarithms to avoid outliers and normalize the variables. All statistical modelling was carried out using Statistical Analysis System Enterprise Guide, version 6.1.

3. Results

3.1. Descriptive analysis

Table 1 indicates 62% (611) of the left behind were female and their ages ranged from 10-14 years (mean age 13.2 years, median 13 years). When compared to non-left behind still majority were female and their mean age was 14.2 years. The fathers for those that were left behind were 97% (953) less likely to be non-farmers. In addition, the majority of left behind young people do not go to school ($p < 0.001$) compared to non-left behind. Left behind young people were more likely staying with grandparents ($p < 0.001$), followed by staying with a guardian 26% (252).

Table 2 shows the cause specific risk of injury for non-left behind and left behind young people. Left behind young people were associated with higher risk

Table 1. Comparison of demographic characteristics between LBAY and non-LBAY

Items	Total, n = 2,991 (%)	Non-Left behind, n = 2,005 (%)	Left Behind, n = 986 (%)	p-values, (χ^2 test)
Gender				
Male	1,318 (44)	943 (47)	375 (38)	0.085
Female	1,673 (56)	1,062 (53)	611 (62)	
Age, Mean		14.2	13.2	
10 -14 years	1,865 (62)	1,303 (65)	562 (57)	0.066
15-24 years	1,126 (38)	702 (35)	424 (43)	
School Attendance				
Yes	2,258 (75)	1,804 (90)	454 (46)	< 0.001
No	733 (25)	201 (10)	532 (54)	
Parental migration				
Internal	906 (92)	-	906 (92)	0.030
International	80 (8)	-	80 (8)	
Father's occupation				
Farmer	72 (3)	39 (2)	33 (3)	0.782
Non Farmer	2,919 (97)	1,966 (98)	953 (97)	
Mother's occupation				
Farmer	39 (1)	21 (1)	18 (2)	0.089
Non farmer	2,952 (99)	1,984 (99)	968 (98)	
Who are staying with				
Aunty/Uncle	101 (3)	-	101 (10)	< 0.001
Grand parents	406 (14)	-	406 (41)	
Parent/s	2,005 (67)	2,005 (100)	-	
Guardian	252 (8)	-	252 (26)	
Alone	227 (8)	-	227 (23)	

**Significant at $p < 0.001$ level. LBAY: Left behind adolescent and youth, non-LBAY: non-left behind adolescent and youth.

Table 2. Injury related emergency department (ED) visits, Hospitalization, and deaths

Items	Non- Left behind n = 2,005 (%)	Left behind n = 986 (%)	Non-Left behind n = 907 (%)	Left behind n = 380 (%)	Non-Left behind	Left behind n = 365 (%)
Poisoning	246 (12.3)	23 (2.3)	80 (8.8)	15 (3.9)	-	6 (1.6)
Fall	1,283 (64.0)	43 (4.4)	659 (72.7)	12 (3.3)	-	-
Drowning	12 (0.6)	8 (0.8)	-	-	-	-
Cut/pierce	163 (8.1)	54 (5.5)	96 (10.6)	19 (5.0)	-	-
Road Traffic	86 (4.3)	391 (39.7)	23 (2.5)	130 (34.2)	-	174 (47.7)
Burns	91 (4.5)	162 (16.4)	38 (4.2)	77 (20.3)	-	2 (0.5)
Suffocation	12 (0.6)	28 (2.8)	-	-	-	-
Machinery	44 (2.2)	13 (1.3)	2 (0.2)	1 (0.3)	-	-
Fire	68 (3.4)	264 (26.8)	9 (1.0)	126 (31.6)	-	183 (50.1)

of injuries especially in road traffic 391(40%), fires 264 (27%), and burns 162 (16%) except for drowning, machinery and poisoning. Moreover, deaths due to injury was highest among left behind (deaths caused by fire was recorded highest with 50.1%, followed by motor vehicle accidents 47.7%) and those left behind recorded the highest in terms of hospitalization cases.

3.2. Bivariate analysis

Table 3 indicates Correlation coefficients (r) to examine strength and significance of linear relationships between the variables in this study.

A significant relationship was found between young people for both cases (left behind and non-left behind). The correlation was considerably lower for the left behind older ($r = 0.780$) than the younger ($r = 0.928$). Parental migration is the most important issue in the total number of injuries. Correlation coefficients show that for non-left behind, there was a greater chance to be classified under poisoning cases ($r = 0.888$) and is

lower compared to the left behind ($r = 0.471$). Again in the cases of falls, the correlation was slightly lower for left behind ($r = 0.332$) than the non-left behind ($r = 0.779$). As expected, road traffic is likely associated with both left behind and non-left behind young people. For the left behind the correlation coefficient shows that it is higher for left behind ($r = 0.934$) and lower among non-left behind ($r = 0.804$). Again in the cases of burns, the correlation was higher for left behind ($r = 0.801$) and lower among non-left behind (0.316). In addition correlation coefficients show that for the left-behind, there is a greater chance for injury caused by fire ($r = 0.925$).

3.3. Multivariate analysis

Table 4 indicates that age (young and old) and gender significantly influence injuries as an outcome. Females have a higher chance of an injury than male, especially those that were left behind, with the odds for gender remaining stable in all three models (OR = 0.25; 95% CI:

Table 3. Correlation between the variables that were examined

Items	LB	NLB	AGE	PSN	FL	DWG	CT	RT	BS	SCN	MCY	FR
LB	1.00											
NLB	0.698	1.00										
AGE	0.928**	0.794*	1.00									
PSN	0.471*	0.888**	0.502	1.00								
FL	0.332*	0.779*	0.230	0.685	1.00							
DWG	0.342	0.638	0.423	0.340	0.591	1.00						
CT	0.425	0.796	0.678	0.272	0.382	0.189	1.00					
RT	0.934**	0.804*	0.512	0.503	0.301	0.221	0.377	1.00				
BS	0.801*	0.316*	0.699	0.239	0.467	0.403	0.509	0.466	1.00			
SCN	0.569	0.211	0.306	0.419	0.288	0.142	0.124	0.521	0.555	1.00		
MCY	0.289	0.380	0.462	0.391	0.640	0.546	0.691	0.399	0.255	0.662	1.00	
FR	0.925**	0.244	0.601	0.360	0.598	0.393	0.193	0.294	0.149	0.501	0.333	1.00

**Significant at $p < 0.01$ level, *Significant at $p < 0.05$ level. Note: LB; Left behind. NLB; Non-left behind. PSN; Poisoning. FL; fall. DWG; Drowning. CT; Cut/pierce. RT; Road traffic. BS; Burns. SCN; Suffocation. MCY; Machinery. FR; Fire.

Table 4. Multivariate linear regression models explaining the injuries

Items	Non-left behind, OR (95% CI)			Left Behind, OR (95% CI)		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
F	0.44 (0.29-0.66)	0.45 (0.29-0.67)	0.43 (0.27-0.65)	0.25 (0.11-0.53)	0.22 (0.10-0.48)	0.21 (0.08-0.46)
YNG	1.81 (1.10-2.98)	1.76 (1.06-2.98)	1.42 (0.88-2.40)	4.61 (1.66-13.10)	3.88 (1.30-11.44)	3.16 (1.06-9.48)
OLD	11.90 (3.30-19.40)	5.20 (3.10-8.60)	6.75(3.60-12.50)	15.60 (14.90-48.80)	10.29 (3.16-33.32)	7.04 (2.11-23.60)
PSN	-	2.38 (1.60-4.12)	2.40 (1.66-4.10)	-	2.72 (1.38-4.18)	2.80 (2.30-4.29)
FL	-	0.82 (0.69-0.99)	0.83 (0.68-1.02)	-	0.74 (0.58-1.00)	0.77 (0.56-1.09)
DWG	-	0.80 (0.68-0.99)	0.88 (0.71-1.10)	-	0.70 (0.54-0.95)	0.78 (0.57-1.06)
CT	-	0.86 (0.72-1.08)	0.84 (0.68-1.12)	-	0.86 (0.68-1.18)	0.82 (0.64-1.12)
RT	-	1.03 (0.88-1.24)	1.05 (0.88-1.25)	-	1.30 (1.08-1.65)	1.30 (1.05-1.62)
BS	-	1.25 (0.83-1.98)	1.27 (0.85-1.90)	-	3.05 (1.68-5.92)	3.20 (1.75-5.88)
SCN	-	1.06 (0.88-1.28)	1.03 (0.88-1.26)	-	1.10 (0.88-1.44)	1.09 (0.84-1.46)
MCY	-	-	1.11 (0.94-1.34)	-	-	1.07 (0.83-1.42)
FR	-	-	5.18 (3.97-8.44)	-	-	6.68 (4.01-11.80)

Note: LB; Left behind. NLB; Non-left behind. YNG; age 10-14. OLD; age 15-24. PSN; Poisoning. FL; fall. DWG; Drowning. CT; Cut/pierce. RT; Road traffic. BS; Burns. SCN; Suffocation. MCY; Machinery. FR; Fire.

0.11-0.53). Moreover, age is significantly associated with increased "injury" status in all models (*i.e.* old has less injuries compared to the young).

When adding types of injuries to model 2 lower levels associations were found for reports of poisoning and falls as an additional risk factor, while falls were only associated with those who are left-behind. Moreover, reports of higher "road traffic" and "burns", turned out to be an important risk factor, for both non- left behind and left behind, except for "cuts/piercing", which revealed no elevated risk for non- left behind.

The final model 3 indicated that fire plays a crucial risk factor in the likelihood of young people, independently of other factors. Left behind young people are highly associated with reports of the fire. Reports on fire had nearly a 7 times higher chance for those left behind, while the odds for non-left behind is only a bit lower (OR = 5.2). In this last model, reports on machinery were significant among non-left behind but not for left behind.

4. Discussion

The migration of parents for jobs or better income to cities has resulted in the separation of many families, leaving their families behind. Most of those left behind are children who cannot move to the city with their parents due to financial constraints (16). China is facing the same situation where over 60 million children in rural areas are left without one or both of their parents (17). The key findings from this study gathered from 2991 young people on the impact of parental migration on injuries in Botswana, supports the hypothesis that parental migration plays a crucial role in injuries of the left behind. Previous studies indicated that the largest cause of death among adolescents and more than seventy percent of these injuries are consequences of motor vehicle collisions (18). This study verifies that absence of parents contributes to road traffic injuries that occur among young people. World Health Organization and the United Nations Children's Fund published the report focused on the five leading causes of child injury deaths being road traffic injuries, drowning, poisoning, burns and falls (19). Recent studies showed that adolescents separated from parents were disadvantaged in social emotional health; for example, and a majority of them have a higher prevalence of depression, suicide, anxiety, and substance abuse than those who are staying with parents. Gender is a significant risk factor for injuries among females in this paper. The previous study reported that being a female has consistently been identified as a significant risk factor for injuries. Therefore, most researchers have suggested that the behavior associated with the female sex role contributes to risk factors for injuries. Parental migration was associated with school dropout among left-behind young people in this study. In the absence of parents, living with grandparents was

associated as a risk factor for injuries. The previous study reported that young people, when both parents migrated, were more likely to be associated with risk factors for injuries. The parents' situation of working in urban areas causes a number of direct, significant effects on the family functionality, especially on the children who remain in rural areas. Due to the problems related to the lack of effective supervision and the absence of a real family environment, children left behind are vulnerable to abuse, labor exploitation and injuries. This category of children has been increasingly affected by lower school performance, with an increasing dropout climax even from early ages (20). Poisoning has been identified as one of the major causes of young people hospital emergencies and admissions in the paper for those that are non-left behind. In African countries, poisoning has also been recognized as a major health problem among adolescents (21). This study reported home burns are among the main causes of morbimortality in Botswana, mainly affecting young people whose parents migrated, and being one of the main accidental reasons for trips to the emergency department. However, WHO highlights a higher frequency of burn injuries in African countries when compared to continents (21). Similar results were obtained in other studies (22), this study observed a greater prevalence of female victims. One possible explanation for this finding may be related to the copied behavior of girls in relation to their mothers since they need to fit in the shoes of the migrated parents. Residential fires are a leading cause of unintentional injury deaths to children in Botswana. In this study, fire injuries accounted for most of the injuries and also registered some deaths in past years. Our data could not distinguish whether the injured adolescent was the same person who started the fire, but other studies have found that as many as one-third of the victims of the fire were the ones who started it (23). Nevertheless, previous studies have shown that injuries are among the top public health problems facing young people today. In the United States in ages 10-19 death from injuries is higher than from all other causes of death (24).

5. Conclusion

This study shows clearly that injuries take a significantly high toll on children's health and left behind young people. Additionally, if parental migration trends continue, Botswana's burden of injuries are expected to rise in the next 10 years. If measures to prevent injuries among the left behind are not taken now, this situation is likely to exacerbate the problem, and not only in Botswana. The Government must take development to rural areas in order to reduce rural- urban migration, and by doing this the numbers of injuries among left behind children may be reduced since children will be under the care of their parents and not grandparents. Young people are among the most vulnerable when it comes to injuries.

The International organizations like United Nations International Children's Emergency Fund must work with partners (governmental and non-governmental) to raise the profile of the preventability of road traffic injuries, falls, fire, burns and promote good practice related to addressing key behavioural risk factors, especially in rural areas because that is an area with a high population of young people who are left behind. Burns and fire are a serious public health problem among the left behind young people in Botswana. The suffering caused by burns is even more tragic as burns are so eminently preventable. High-income countries have made considerable progress in lowering rates of burn deaths, through a combination of proven prevention strategies and through improvements in the care of burn victims. These can be a learning tool for Botswana and this would likely lead to significant reductions in rates of burn-related death and disability among left behind young people.

Acknowledgements

United Nation Economic Commission for Africa (Social Development Policy Division) supported this work without a grant. We thank Ministry of Health for providing data on hospitalization and emergency department visits, and the department of civil and national registration on death causes by injuries among people aged 10 years to 24 years. In addition, I would like to thank William Muhwava for support and guidance.

References

- Gibson J, McKenzie D, Stillman S. The impacts of international migration on remaining household members: Omnibus results from a migration lottery program. *The Review of Economics and Statistics*, 2011; 93:1297-1318.
- Mu R, Van de Walle D. Left behind to farm? Women's labor reallocation in rural China. *Labour Economics*. 2011; 18(S1):S83-S97.
- United Nations International Children's Fund. *Migration and Youth: Challenges and Opportunities*, 2014. <http://unesdoc.unesco.org/images/0022/002277/227720e.pdf> (accessed February 14, 2018).
- Cortina J. Beyond the Money: The Impact of international migration on children's life satisfaction: Evidence from Ecuador and Albania. *Migration and Development*. 2014; 3:1-19
- Jones SE, Shults RA. Trends and subgroup differences in transportation-related injury risk and safety behaviors among US high school students, 1991-2007. *J Sch Health*. 2009; 79:169-176.
- World Health Organization, *World report on disability*. Geneva, Switzerland, 2011. http://www.who.int/disabilities/world_report/2011/report.pdf (accessed February 20, 2018)
- Antman FM. Gender, educational attainment, and the impact of parental migration on children left behind. *J Popul Econ*. 2012; 25:1187-1214.
- United Nations Development Programme. 2016 Human Development Report" http://hdr.undp.org/sites/default/files/2016_human_development_report.pdf (accessed March 2, 2018)
- Botswana or Batswana? It's complicated - Voices of Africa". *Voices of Africa*. <http://voicesofafrica.co.za/botswanan-batswana-its-complicated/> (accessed January 16, 2018).
- Campbell EK, Oucho JO. (2003), *Changing attitudes to immigration and refugee policy in Botswana* (PDF), Migration Policy Series, 28, Cape Town, South Africa/Toronto, Canada: Southern African Migration Project. 2003. ISBN 1-919798-47-1
- Denbow J, Thebe PC. (2006). *Culture and Customs of Botswana*. Westport, CT: Greenwood Press. London, 2006; 2-12.
- Kwak E, Rudmin F. Adolescent health and adaptation in Canada: Examination of gender and age aspects of the healthy immigrant effect. *Int J Equity Health*. 2014; 13:103.
- Zhou FL, Duan CR: Literature review on studies on left behind children. *Population Research*. 2006; 3:60-65.
- International Monetary fund, Botswana. IMF Country report, 2017 <http://www.imf.org/en/Publications/CR/Issues/2017/08/08/Botswana-2017-Article-IV-Consultation-Press-Release-Staff-Report-45172> (Accessed January 20, 2018).
- Government of Botswana. UNDP Botswana Annual Report 2013/2014. <http://www.bw.undp.org/content/botswana/en/home/library/poverty/undp-botswana-annual-report-2013-2014.html> (accessed February 14, 2018).
- Shen M, Wang Y, Yang S, Du Y, Xiang H, Stallones L. Agricultural exposures and farm-related injuries among adolescents in rural China. *Inj Prev*. 2013; 19:214-217.
- Li S. The economic situation of rural migrant workers in China. *China Perspectives*. 2010; 4:4-15.
- Wen M. The effect of family structure on children's health and well-being: Evidence from the 1999 National Survey of America's Families. *Journal of Family Issues*. 2008; 29:1492-1519.
- Zhou C, Sylvia S, Zhang L, Luo R, Yi H, Liu C, Shi Y, Loyalka P, Chu J, Medina A, Rozelle S. China's left-behind children: Impact of parental migration on health, nutrition, and educational outcomes. *Health Aff (Millwood)*. 2015; 34:1964-1971.
- Hathaway J. Forced migration studies: Could we agree just to 'date'? *J Refug Stud*. 2007; 20:349-369.
- World Health Organization. *Health for the world's adolescents. A second chance in the second decade*. <http://apps.who.int/adolescent/second-decade> (Accessed January 9, 2018).
- Saunders NR, To T, Parkin PC, Guttman A. Emergency department revisits by urban immigrant children in Canada: A population-based cohort Study. *J Pediatr*. 2016; 170:218-226.
- Laursen B, Moller H. Unintentional injuries in children of Danish and foreign born mothers. *Scand J Public Health*. 2009; 37:577-583.
- United Nations Children' Fund. A league table of children deaths by injury in rich nations. *Innocenti Report Card*. 2001; Issue No.2. <https://www.unicef-irc.org/publications/pdf/repcard2e.pdf> (accessed February 2, 2018).

(Received April 2, 2018; Revised April 24, 2018; Accepted April 26, 2018)

Internal migration and regional differences of population aging: An empirical study of 287 cities in China

Rong Chen^{1,*}, Ping Xu², Fen Li¹, Peipei Song^{1,3}

¹Shanghai Health Development Research Center, Shanghai, China;

²Department of Gerontology, University of Massachusetts Boston, MA, USA;

³Graduate School of Frontier Science, The University of Tokyo, Kashiwa-shi, Chiba, Japan.

Summary

In addition to birth and death, migration is also an important factor that determines the level of population aging in different regions, especially under the current context of low fertility and low mortality in China. Drawing upon data from the fifth and sixth national population census of 287 prefecture-level cities in China, this study explored the spatial patterns of population aging and its trends from 2000 to 2010 in China. We further examined how the large-scale internal migration was related to the spatial differences and the changes of aging by using multivariate quantitative models. Findings showed that the percentage of elder cities (*i.e.* proportion of individuals aged 65 and above to total population is higher than 7%) increased from 50% to 90% in the total 287 cities within the decade. We also found that regional imbalances of population aging have changed since 2000 in China. The gap of aging level between East zone and the other three zones (*i.e.* West, Central, and North-east) has considerably narrowed down. In 2000, Eastern region had the greatest number (65) of and the largest proportion (74.7%) of elder cities among all four regions. By 2010, the proportion (87.4%) of elder cities in the eastern region was slightly lower than Central (91.4%), Western (88.2%) and North-east sectors (91.2%). Results from multivariate quantitative models showed that the regional differences of population aging appear to be affected much more by the large-scale internal migration with clear age selectivity and orientation preference than by the impact of fertility and mortality. Population aging is expected to continue in China, which will in turn exacerbate regional imbalances. Policies and implications are discussed to face the challenges that the divergent aging population may present in China.

Keywords: Regional differences of aging, internal migration, multivariate quantitative models

1. Introduction

Aging is an inevitable path of demographic transition, as the inescapable result of low fertility and long life (1). China has experienced a population transition in a relatively short period, which can be divided into three phases: mortality-dominated population dynamic stage (1949-1970), fertility-dominated population dynamic stage (1971-1999) (2), and accelerating trend of aging

stage (since 2000). After the establishment of the new government in China in 1949, medical and sanitation conditions were greatly improved, resulting in a huge mortality rate declination in China (3). Concurrent with Chinese economic reforms in 1978, China has experienced profound social and economic development (4). In addition, due to implementation of the family planning policy in China at the end of 1970s, China's total fertility rate decreased from 6 in 1970 to 2.1 in 1991, and to 1.7 in 2000 (5). Remaining low rates of fertility and mortality, and increased longevity has changed the age structure of the Chinese population. China's population has been aging at unprecedented levels. In 2000, the proportion of older adults aged 65 and over reached 7% (6), indicating that China has entered into an aging society (defined as a society with 7-14% of the total population aged 65 and over), and the

Released online in J-STAGE as advance publication April 2, 2018.

*Address correspondence to:

Dr. Rong Chen, Shanghai Health Development Research Center, Room707 No. 122 South Shanxi Road, Shanghai 200040, China.

E-mail: chenrong8721@126.com

proportion of older population in the total population has increased to 10.8% in 2016 in China (6). Furthermore, it is predicted that China will become a super-aged country (a society with 20% or more of the total population aged 65 and over) in less than 20 years (7). The outpaced growth of aging over that of economic development and the social security system will put tremendous pressure on older individuals, their families, as well as on society (8-10).

In addition to low fertility rates and low mortality rates, population migration is the third major factor that determines population age structure. Accompanied with the process of urbanization and modernization in China during three decades after economic reforms, lessened restrictions for residents moving from one place to another resulted in a rapidly increased number of temporary internal migrations (also called as "floating population") (11). There were about 121 million floating populations in 2000. This number increased to 221 million in 2010, and reached to 245 million in 2016 (accounting for 18% of the country's population) (6). Internal migration often happened among young and middle-age adults (12), and thus exerted significant effects on age structure for both out-migration and in-migration areas (13). Specifically, recipient places of migration postponed or slowed down the aging process, while sending places of migration exacerbated the aging process (3). As changes in fertility rates and mortality rates have remained in a more stable stage, large-scale internal migration has been playing an increasingly important role in altering the regional demographic structure patterns and shaping the regional aging trends (14-15). Previous studies have found that internal migration have very opposite effects on urban and rural population aging: diluting the degree of aging and slowing down the speed of aging in urban areas, while speeding up aging process and exacerbating the degree of aging in rural areas, resulting in inversion of urban and rural aging (16-18).

Previous research has explored spatial patterns of aging before 2000 in China. Wu and colleagues found that the degree of aging was consistent with the level of economic development in different regions, and the differences among regions were considerably large (19). Relatively fewer studies have, however, focused on spatial patterns of aging and its changes in more recent years. It is remarkably informative in the present study to provide more comprehensive and dynamic pictures by showing the spatial patterns of aging both in 2000 and in 2010 in China. In addition, most prior studies on regional aging trends divided the whole country into urban and rural areas (17,18), or three regions, *i.e.* eastern, middle, and western region (19), or at the level of 31 provinces (20). One of the limitations of previous divisions is that the spatial scope is relatively large at the rural/urban level and/or province level, as the regional aging degree may be averaged. When the spatial scale becomes smaller, the

regional differences and changes of population aging will be more obvious and sharp. However, if the spatial scale is too small (for example, at the county level and below), although it can be a typical case study, it may lack generalization. Therefore, we employed our analyses at the level of prefecture-level cities.

This paper aimed to examine the new characteristics and trends of regional imbalance in aging at the level of prefecture level cities in China during the period of time from 2000-2010, and to explore how the massive internal migration as well as other demographic factors were related to spatial distributions and its trends of the aging process, utilizing data from the fifth and sixth national population census (2000 and 2010 census).

2. Materials and Methods

2.1. Data source

Population information of the 287 cities was from tabulations in 2000 census and 2010 census of China by county (21,22). Population data resources of the 31 provinces were from tabulations in 2000 and 2010 census of China (23,24). In order to accurately measure the role of population migration, we utilized residential population (*i.e.* the de facto population), rather than the *hukou* population (*i.e.* the de jure population). Because the internal migration population in China, also called floating population, moved to the receiving places, but their *hukou* were still in the sending places. From 2000 to 2010 census, the administrative divisions of some prefecture-level cities were adjusted. Accordingly, we modified the administrative divisions of those prefecture-level cities year by year, following the rules provided by the Ministry of civil affairs of China.

The unit of analysis in this study was 287 cities from 31 provincial units in China, including 23 provinces, 5 autonomous regions and 4 municipalities directly under the central government in mainland China. We assigned the 31 provincial units into four regions, namely East, Central, West and North-east. East included 10 provincial units: *Beijing, Tianjin, Hebei, Shanghai, Jiangsu, Zhejiang, Fujian, Shandong, Guangdong* and *Hainan* provinces. Six provinces of *Shanxi, Anhui, Jiangxi, Henan, Hubei* and *Hunan* provinces were in the Central region. West had 12 provincial units of *Inner Mongolia, Guangxi, Chongqing, Sichuan, Guizhou, Yunnan, Tibet, Shaanxi, Gansu, Qinghai, Ningxia* and *Xinjiang*. Lastly, three provinces, *Liaoning, Jilin* and *Heilongjiang*, were in the North-east region (Figure 1). The 287 cities in the paper are actually prefecture-level or higher cities, including four municipalities directly under the central government and 283 prefecture-level cities, excluding the other 50 prefecture-level administrative units. According to the administrative level, the 287 cities were also dichotomized into core city and non-core city. Core cities included 4 directly administrated municipalities,

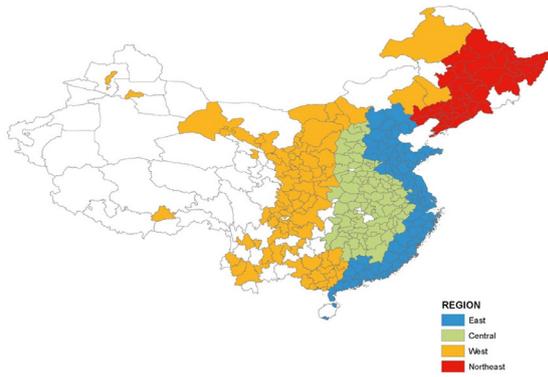


Figure 1. Cities in four regions, China, 2010.

and 31 specific plan oriented cities, sub-provincial cities, and provincial capitals. The rest of the 252 prefecture-level cities were non-core cities.

2.2. Method

2.2.1. Dependent variables

The dependent variables included two aspects of aging across 287 cities in China: the level of aging and the rate of increase in aging between 2000 and 2010. The level of aging for each city was measured by the ratio of the number of older adults aged 65+ to the population as a whole (for convenience, hereafter referred to "PA"). Based on the ratio, we categorized the 287 cities into three groups: *i*) young city where the ratio was lower than or equal to 4% (*i.e.* $PA \leq 4\%$); *ii*) adult city where the ratio was greater than 4% but lower than or equal to 7% (*i.e.* $4\% < PA \leq 7\%$); and *iii*) elder city where the ratio was greater than 7% (*i.e.* $PA > 7\%$). We further subdivided elder cities into two categories: young old city where the ratio was between 7% and 10% (*i.e.* $7\% < PA \leq 10\%$) and old old city where the ratio was higher than 10% (*i.e.* $PA > 10\%$).

The rate of increase in aging was the change in the level of aging from 2000 census to 2010 census (*i.e.* PA in 2010 subtracted by PA in 2000, hereafter referred to "ΔPA"). For the purpose of descriptive statistics, the 287 cities were assigned into 5 categories: *i*) decrease in aging (*i.e.* $\Delta PA < 0$); *ii*) low rate of increasing (*i.e.* $0 \leq \Delta PA \leq 1\%$); *iii*) medium rate of increasing (*i.e.* $1\% < \Delta PA \leq 2\%$); *iv*) medium-to-high rate of increasing (*i.e.* $2\% < \Delta PA \leq 3\%$); and *v*) high rate of increasing (*i.e.* $\Delta PA > 3\%$). In addition, for further exploratory analysis, continuous ΔPA was utilized in regression models. In the statistical equations, ΔPA was denoted as *agingchange_i*.

2.2.2. Key independent variable

The rate of net in-migration population was the average value of rates across the 287 cities in the time of 2000 census and 2010 census, denoted as *nmpri_i*. The rate of

net in-migration population was equal to the number of net in-migration population divided by the number of the residential population in each city. Initially, the number of net in-migration population of the cities was calculated by the number of in-migration population minus the number of out-migration population. Due to data limitation, the number of in-migration population and out-migration population were not available at prefecture-level cities in the tabulations in 2000 census and 2010 census of China by county. Therefore, we captured the number of net in-migration population by the number of residential population subtracted by the number of *hukou* population. Although this calculation may cause discrepancies because of a small amount of unregistered people, the variance is minor.

2.2.3. Covariates

We also controlled for a set of relevant variables. The natural population growth rate was calculated by the average value of resident population in 2000 and 2010, indicated as *npgr_i*. The natural growth rate of residential population was captured by subtracting the number of population migration growth from population growth and divided by the number of resident population. In addition, the rate of population aging at 2000 was controlled, denoted as *agingr_i*. Furthermore, in order to explore regional differences, we assigned regions as 4 categorical dummy variables, with East as reference group, and the rest denoted as *northeast_i*, *central_i*, and *west_i*. The last control variable was a dummy variable, whether or not a core city, with 1 = yes, 0 = no, denoted as *corecity_i*.

2.2.4. Analytical strategy

We first presented maps to show spatial description of the level of aging in both 2000 and 2010 as well as the changes from 2000 and 2010. Maps were generated in ArcGIS 10 software. The chi-square test was carried out to compare the differences in the degree and speed of aging among the four regions. We further explored the association between population migration and the rate of increase in aging, using multivariate Ordinary Least Squares (OLS) regression models. Four models were examined step by step (see equations below). To be specific, Model 1 only included the rate of residential population growth. Model 2 added the rate of net in-migration population. Model 3 incorporated the level of aging at baseline. Model 4 was a full model, adding into regional variables and a core city variable. Analyses were conducted in SPSS version 19.

$$agingchange_i = \alpha_1 + \alpha_2 npgr_i + u_{1i} \tag{1}$$

$$agingchange_i = \beta_1 + \beta_2 nmpri_i + \beta_3 npgr_i + u_{2i} \tag{2}$$

$$agingchange_i = \gamma_1 + \gamma_2 nmpri_i + \gamma_3 npgr_i + \gamma_4 agingr_i + u_{3i} \tag{3}$$

$$\begin{aligned}
 \text{agingchange}_i &= \delta_1 + \delta_2 \text{nmpr}_i + \delta_3 \text{npgr}_i + \delta_4 \text{agingr}_i \\
 &+ \delta_5 \text{northeast}_i + \delta_6 \text{central}_i + \delta_7 \text{west}_i \\
 &+ \delta_8 \text{corecity}_i + u_{4i} \quad (4)
 \end{aligned}$$

3. Results

3.1. Descriptive results on regional differences of population aging

Table 1 and Figure 2 show the level of aging for 287 cities in China in 2000 in the form of table and map. In

general, only 49.8% of cities were elder cities in 2000, however, this number increased to 89.2% in 2010. Most of cities entered into elder cities within a decade. Although China was, on average, considered as an aging society in 2000, each city had a different pace. In the 2000 census, about 143 cities (49.8%) belonged to elder cities, among which 4 cities' PA were higher than 10%, including *Shanghai* (11.5%), *Nantong* (12.4%) and *Taizhou* (10.3%) in Jiangsu province, and *Lishui* (10.0%) in Zhejiang province. In addition, 139 cities (48.4%) were in the adult city group. Only 5 cities were young

Table 1. The level of aging for 287 cities in China in 2000 census

Items	Young city	Adult city	Elder city		Total
			Young old city	Old old city	
Number of East cities	2	20	61	4	87
% of East cities	2.3	23.0	70.1	4.6	100
Number of Central cities	0	46	35	0	81
% of Central cities	0	56.8	43.2	0	100
Number of West cities	3	50	32	0	85
% of West cities	3.5	58.8	37.6	0	100
Number of North-east cities	0	23	11	0	34
% of North-east cities	0	67.6	32.4	0	100
Number of above cities	5	139	139	4	287
% of above cities	1.7	48.4	48.4	1.4	100

Note: Pearson $\chi^2 = 42.954, p = 0.000$. Chi square tested regional differences on the level of aging in 2000, showing that there were significant differences among the four regions.

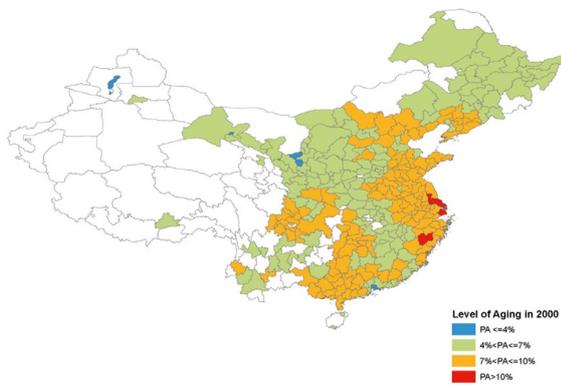


Figure 2. The level of aging for 287 cities in China in 2000 census.

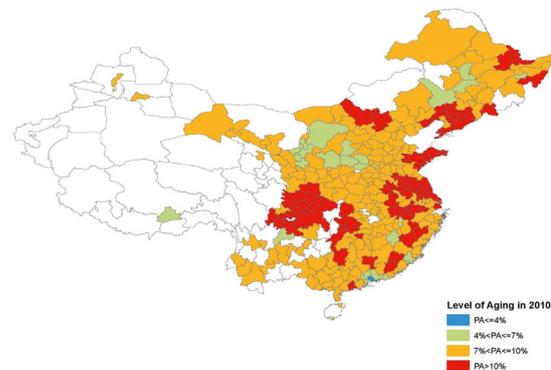


Figure 3. The level of aging for 287 cities in China in 2010 census.

Table 2. The level of aging for 287 cities in China in 2010 census

Items	Young city	Adult city	Elder city		Total
			Young old city	Old old city	
Number of East cities	2	9	50	26	87
% of East cities	2.3	10.3	57.5	29.9	100
Number of Central cities	0	7	55	19	81
% of Central cities	0	8.6	67.9	23.5	100
Number of West cities	0	10	55	20	85
% of West cities	0	11.8	64.7	23.5	100
Number of North-east cities	0	3	17	14	34
% of North-east cities	0	8.8	50.0	41.2	100
Number of above cities	2	29	177	79	287
% of above cities	0.7	10.1	61.7	27.5	100

Note: Pearson $\chi^2 = 14.611, p = 0.263$. Chi square tested regional differences on the level of aging in 2010, showing that there were no significant differences among the four regions.

Table 3. Changes on Level of Aging for 287 Cities in China, from 2000 to 2010

Items	Decrease in aging	Low rate of increasing	Medium rate of increasing	Medium-to-high rate of increasing	High rate of increasing	Total
Number of East cities	6	30	30	16	5	87
% of East cities	6.9	34.5	34.5	18.4	5.7	100
Number of Central cities	0	12	33	25	11	81
% of Central cities	0.0	14.8	40.7	30.9	13.6	100
Number of West cities	1	2	26	23	33	85
% of West cities	1.2	2.4	30.6	27.1	38.8	100
Number of North-east cities	0	0	5	21	8	34
% of North-east cities	0.0	0.0	14.7	61.8	23.5	100
Number of above cities	7	44	94	85	57	287
% of above cities	2.4	15.3	32.8	29.6	19.9	100

Note: Pearson $\chi^2 = 92.739, p = 0.000$. Chi square tested regional differences on the rate of increase in aging from 2000 census to 2010 census, showing that there were significant differences among the four regions.

cities, which were *Shenzhen* (1.2%) and *Dongguan* (2.1%) in Guangdong province, *Jiayuguang* (3.8%) in Gansu province, *Zhongwei* (4.0%) in Ningxia province, and *Kelamayi* (3.8%) in Xinjiang province. With regard to regional differences in 2000, East area had the greatest number of cities which belonged to elder cities (taking up about 74.7%), followed by Central area (43.2%). West area had about 37.6% of elder cities and North-east had the least proportion of elder cities (32.4%). More importantly, the regional differences were statistically significant in 2000 ($\chi^2 = 42.954, p = 0.009$), which meant that the level of aging in East was significantly higher than that of other geographic areas.

Table 2 and Figure 3 show the level of aging across China in 2010. As shown, only two cities, *Shenzhen* and *Dongguan*, were young cities (1.8% and 2.3%, respectively). The number of adult cities reduced from 139 in 2000 to 29 in 2010. In total, only 31 cities (about 10.8%) were not elder cities across China in 2010. In comparison, the number of old old cities increased from 4 in 2000 to 79 in 2010. Specifically, *Nantong* and *Taizhou* ranked as the top 1 and 2 among the old old cities, both of whose PA increased by 4% from 2000 to 2010. The portion of senior population in *Shanghai* decreased from 11.5% in 2000 to 10.1% in 2010, indicating a slowed down pace in aging.

Regional differences in the level of aging in 2010 are also presented in Figure 3. As depicted, a notable number of cities in the Central, West and North-east areas entered into elder cities in 2010, about 91.4%, 88.2%, and 91.2%, respectively. In particular, among the 34 cities in the North-east region, about 41.2% of cities' PA were greater than 10%, which was higher than those in other regions. On the contrary, fewer cities in the East region became elder cities. The regional differences were not statistically significant in 2010 ($\chi^2 = 14.611, p = 0.263$), indicating that the spatial gap became smaller and smaller in terms of the rate of aging, since almost all cities entered into elder cities.

Table 3 and Figure 4 present the rate of increase in aging from 2000 census to 2010 census. The three

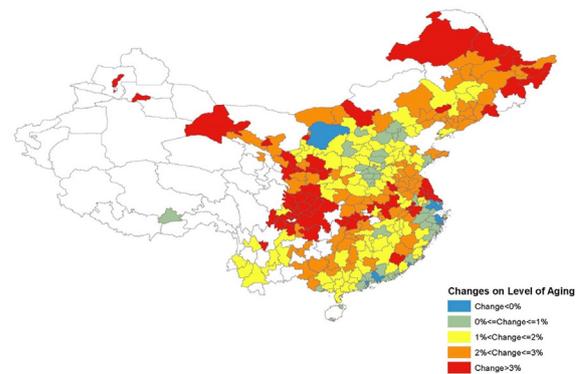


Figure 4. Changes on level of aging for 287 cities in China, from 2000 to 2010.

regions, North-east, Central, and West, had greater rates of increase in aging than the East area. And the spatial differences were statistically significant ($\chi^2 = 92.739, p = 0.000$). In particular, the North-east area had the fastest pace of population aging, followed by West area and Central area. East region had the lowest rate of increase in aging. It is worth noting that 6 cities had negative rate of increase in the level of aging, including *Shanghai*, *Suzhou*, *Xiamen*, *Huizhou*, *Zhongshan*, *Eerduosi*, and *Ningbo*.

3.2. Domestic population migration

Spatial disparities in economic development have been a historical and outstanding issue in China, with East and coastal areas substantially advanced compared to other regions due to location advantages, market forces, and favorable policies. For instance, three metropolitan areas, namely triangle cities along Zhujiang, triangle cities along Yangzi River, and Jing-jin-qi cities have the most economic growth. Motivated by more job opportunities and better earnings in economically developed areas, a large-scale of working-age population internally migrate from rural areas to urban areas, from less developed cities to more developed cities, and from West and

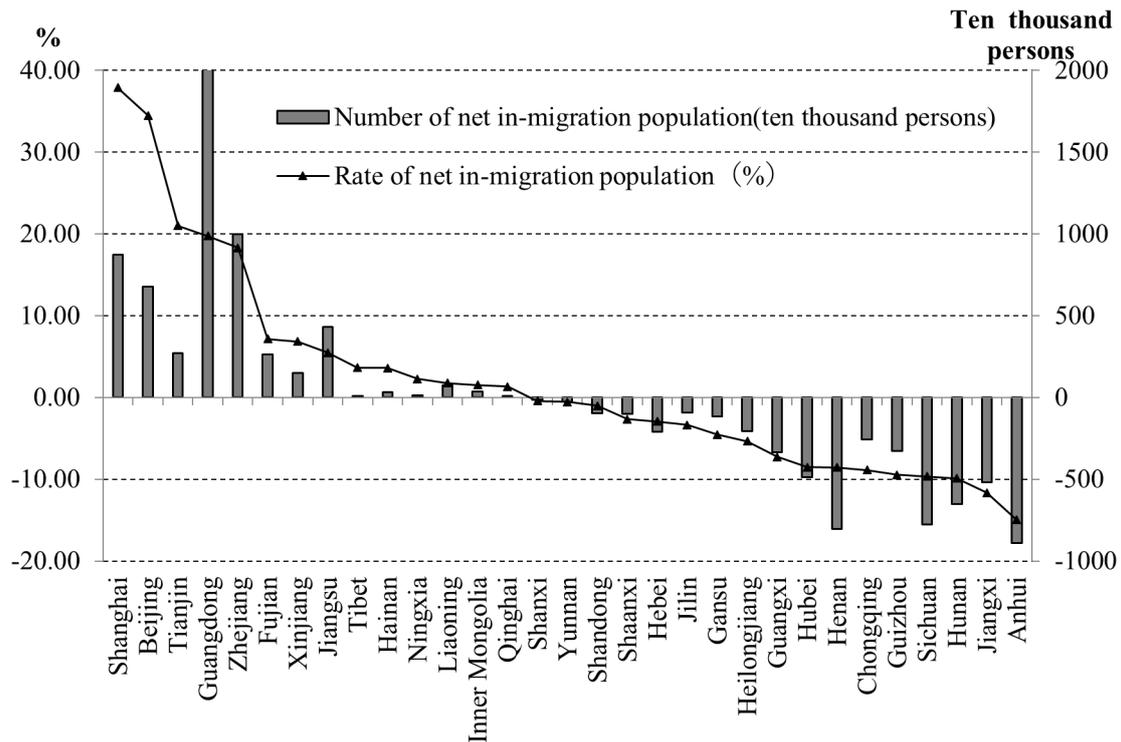


Figure 5. The number and rate of net in-migration population of 31 provinces in 2010 census. Note: The number of net in-migration population of the provinces calculated by the number of in-migration population minus the number of out-migration population, and the rate of net in-migration population was equal to the number of net in-migration population divided by the number of the residential population in each province.

Central areas to East areas.

From the perspective of provinces, in 2010 census, 14 provinces belonged to in-migration areas, most of which are located in the East; whereas 17 provinces belonged to out-migration areas, most of which are located in the Western and Central China. For instance, 8 out of 10 provinces in East region were in-migration provinces, particularly Guangdong province had the largest size of in-migration population, reaching to 20 million. Furthermore, Shanghai, Beijing, and Tianjin were the three cities with highest rate of in-migration population, about 37.9%, 34.5%, and 21%, respectively. On the contrary, all the six provinces in the Central region belonged to out-migration provinces; in particular Henan province and Anhui province had the greatest number of out-migration population, about 8 million and 8.9 million, respectively. Moreover, 5 out of 11 provinces in West area were in-migration provinces, with a relatively small scale of migration population though. Xinjiang province had the highest rate of in-migration, whereas Sichuan province had the greatest rate of out-migration in the Western China. Lastly, within the three provinces in the North-east area, Heilongjiang and Jilin belonged to in-migration provinces. Liaoning was an out-migration province, although both the size and rate of in-migration were low (Figure 5).

Corresponding to internal migration patterns at the province level, migration population were also

concentrated to certain cities, such as capital cities, and hub cities in those three advanced metropolitan areas defined above. The number of in-migration cities decreased from 125 in 2000 to 103 in 2010, while the number of out-migration cities increased from 162 in 2000 to 184 in 2010. Loss of net population cities were mainly located in areas with agriculture and underdeveloped economic growth. During 2000 to 2010, the proportion of the 35 core cities in the national population increased from 19.8% to 22.3%, while non-core cities' percentage of population fell from 73.2% to 70.9%, reflecting population agglomeration to the core cities. Therefore, the internal massive migration has obvious directional preferences, due to regional unbalanced economic development.

In addition, internal migration in China has distinct age selectivity as well. The 2010 census showed that 65% of migration population was the aged 20-49 adult labor force, specifically about 27.2% was between 20-29 years old; and 21.3% was between 30-39 years old, and the rest 16% was between 40-49 years old. In comparison, in the whole population, about 17.1% aged 20-29, 16.1% aged 30-39, and 17.3% aged 40-49, showing that the structure of floating population was younger. Large-scale of internal migration with obvious directional preferences and age selectivity will alter the prevalence of older population in both in-migration and out-migration cities.

Table 4. The regression results from OLS models

Items	Model 1	unadjusted Model 2	adjusted Model 2	Model 3	Model 4
$npgr_i$	-0.101** (0.023)	-0.142** (0.021)	-0.347**	-0.232** (0.021)	-0.224** (0.023)
$nmpri_i$		-0.040** (0.004)	-0.480**	-0.055** (0.004)	-0.042** (0.005)
$agingr_i$				-0.370** (0.040)	-0.229** (0.047)
$northeast_i$					0.345* (0.197)
$central_i$					0.345** (0.133)
$west_i$					0.996** (0.144)
$corecity_i$					-0.298* (0.165)
c	2.569** (0.136)	-2.730** (0.121)		5.727** (0.344)	4.336** (0.448)
N	287	287	287	287	287
adj.R2	0.058	0.277	0.277	0.441	0.524

Note: * $0.05 < p < 0.1$; ** $p < 0.01$. Numbers in parentheses are standard deviations.

3.3. OLS regression results

Table 4 shows the regression results from OLS models. As depicted in Model 1, rate of natural population growth was negatively associated with the rate of increase in aging population ($b = -0.101$, $p < 0.001$). Unadjusted and adjusted Model 2 showed that rate of net in-migration population was negatively associated with increasing rate of aging (*adj.* $b = -0.480$, $p < 0.001$). It indicates that both the increase (or decrease) of natural growth rate and net in-migration rate explained certain variances on the dependent variable, diluting (or exacerbating) the level of population aging. We noticed that the absolute coefficient of in-migration population rate was larger than that of residential population growth rate in adjusted Model 2, indicating population migration had a greater relationship with population aging than those with fertility and mortality. Jones has also pointed out that regional differences in population aging appear to be affected much more by patterns of internal migration than by differences in fertility and mortality (25).

Furthermore, Model 3 showed a significantly negative association between baseline rate of aging and the increasing rate of aging ($b = -0.370$, $p < 0.001$). That is, cities with a high level of aging in 2000 had a relatively slower increase in the level of aging in 2010, whereas those with low level of aging in 2000 had a comparatively faster rate of increase in aging in 2010. The results indicated a smaller gap among all cities in terms of level of older population, during the decade between 2000 and 2010. The rate of in-migration population remained significant in Model 3. Lastly, results in Model 4 demonstrated that compared to cities in the East area, those in West, Central, and North-east areas had a significantly higher rate of

aging increase. Specially, West area had highest rate of increase in aging ($b = 0.996$, $p < 0.01$), followed by Central and North-east areas (both $b = 0.345$). Core cities had a lower rate of aging increase compared to non-core cities ($b = -0.298$, $0.05 < p < 0.1$), indicating that administratively smaller cities had a faster pace of aging than those of hierarchically higher rank cities.

In conclusion, these results suggested that internal migration in China has noticeable patterns for directional preferences and age selectivity, which ultimately impacts the prevalence and incidence of population aging for every province and city. Migration of working-age adult labor force diluted the in-migration cities' (mostly in East area and core cities) level of aging and its rate of increase in aging population; whereas by contrast it increased the out-migration cities' (mainly in West and Central, agricultural areas) level of aging.

3.4. Representative cities

As stated earlier, *Shenzhen* and *Dongguan* were the two youngest cities in both 2000 and 2010, indicated by no increase in the level of aging within 10 years. Also, *Beijing*, *Tianjin* and *Guangzhou* had a lower rate than national average rate of increase in aging. Further, *Shanghai* had a negative rate of increase in aging (-1.34%). These cities shared a common reason, that is, a large scale of young labor force migrated into these places, delaying or reversing the speed of population aging. For instance, the size of in-migration population in *Shenzhen* was 5 times larger than its *hukou* population, similarly about 4 times bigger in *Dongguan*.

Take *Shanghai* for another example. The 2000 census and 2010 census showed that, in-migration population in Shanghai aged 15-59 took up 88.9% and

85.6%, respectively, among which labor force aged 20-44 made up about 70.3% and 69.1%, respectively. In 2010, the median age of *hukou* population in Shanghai was 44.1, and the proportion of older adults aged 65 and above was 15.8%. Due to the large scale of young and middle-aged labor forces moving in, the median age of resident population decreased to 37.8 years old and the percentage of older population was reduced to 10.1% (26). A similar situation occurred in other core cities, such as *Suzhou*, *Zhongshan*, and *Xiamen*.

There is also a compelling story for 13 cities in *Jiangsu* province, which can be divided into southern part and northern part. From 2000 to 2010, cities in the northern part had a negative growth of population, while most southern cities' population growth was positive, and the rates of increase in aging in the northern cities were significantly higher than that in southern cities, though fertility rate in the vast areas of the northern part has been relatively higher than that in the south (27). It is plausible that the cities in northern part of the province are economically less developed than the southern cities. A large amount of adult labor force moves out from the northern part, which ultimately results in population loss and rapidly aging process in most cities of the northern part. On the contrary, the pace of aging in these southern cities has been slowed down due to massive immigration.

4. Discussion

4.1. The pattern of regional imbalance in aging has changed within a decade in China

A population with a low fertility rate is likely to have a high proportion of older adults and vice versa (3). The global demographic transition over the past three centuries has been summarized as: the transition began with mortality decline, followed by fertility decline, and ended up with an aging population (1). At the national level, China has stepped into an aging society since 2000 and has still been experiencing a progressive process of aging, as a result of the rapid decline in fertility and persistent longevity. What is more noteworthy is that the regional aging patterns and trends vary a lot within China due to imbalanced social economic development and population transitions. This paper examined such regional differences and the changes in 287 cities from the degree of and speed of aging from 2000 to 2010.

Results show that the percentage of cities, which have entered the aging society increased from 49.8% to 89.2% in the total of 287 cities within a decade (Table 1 and Table 2), indicating the accelerated process of population aging in the country. We also found that the regional imbalance of population aging has changed in China since 2000. In 2000, Eastern region had the most number (65, Table 1) of and the largest

proportion (74.7%, Table 1) of elder cities among all four regions. However, cities in Central, Western and North-east sectors had a faster pace in aging from 2000 to 2010, significantly faster than those cities in the Eastern sector. Also, the number of new elder cities was far more than that in the Eastern zone. By 2010, the proportion (87.4%, Table 2) of aging cities in the eastern region was even slightly lower than Central (91.4%, Table 2), Western (88.2%, Table 2) and North-east sectors (91.2%, Table 2). The gap of aging level between the other three zones and the Eastern zone has considerably narrowed down.

These findings have revealed dynamic relationships between economic development and population aging. In 2000, the aging level was much higher in more developed areas than that in less developed ones. However, the pattern has changed within a decade. The less developed areas (e.g. cities located in the north of *Jiangsu* province) had a faster pace of population aging from 2000 to 2010 than those of more advanced areas. We also found that the rate of increase in aging has slowed down in certain advanced areas, and the level of aging even declined in several cities, such as *Beijing*, *Shanghai*, *Shenzhen*, *Dongguan*, *Xiamen*, *Suzhou*, and *Nanjing*.

4.2. The large-scale internal migration with obvious age selectivity and orientation preference has played a critical role in reshaping spatial patterns of population aging in China

There is about 245 million of floating population in China, accounting for 18% of the country's whole population. The large-scale internal migration has become the most prominent population phenomenon in China's economic and social development. Gu (2014) has suggested that internal migration might have dominated population dynamics in China, and called more attention to the impacts of population movement (27). The results of multivariate quantitative models showed that the regional differences of population aging appear to be affected much more by the massive internal migration than by differences in fertility and mortality. It is possible that the massive internal migration since the late 1980s has been showing an apparent orientation preference, often from rural areas to urban areas, from West and Central regions to East region, and from economically less developed cities to more advanced ones. In addition, migration often happened among young and middle-age adults, exerting contradictory effects on age structure for out-migration and in-migration areas: diluting the degree of aging and slowing down the speed of aging in recipient places of migration, while speeding up aging process and exacerbating the degree of aging sending places of migration.

4.3 How to deal with the challenges of regional aging

and the massive internal migration has been an emerging and urgent issue for research scholars and policy makers

The massive internal migration was partially due to unbalanced regional economic development, leading to unequal level of income and access to employment opportunities. Due to a large amount of outflow labor force, some areas in Central and Western China have been facing a series of problems in economic development, such as reduction of labor productivity and total consumption, shrinkage of agriculture industry, and so forth. In addition, the massive internal migration and regional aging trends were bound to exacerbate the regional imbalance of pension funds. It is estimated that the inter-provincial floating population paid more than 50 billion *yuan* pension in the receiving regions in 2010. East region received more than 30 billion *yuan*, while the Central and West regions lost about 24 billion *yuan* and 8 billion 400 million *yuan* (28). For the less developed areas, pension coverages are relatively scarce, and the burden will become more and more heavy, which will inevitably hinder their economic growth.

Furthermore, internal migration in China will continue in the future. The less developed areas may encounter various challenges, as its population size is getting smaller while the portion of older adults is getting larger. Due to population reduction, aging, serious loss of labor, infrastructure lack of economies of scale, gradual depression and even abandoned phenomenon may occur in some villages, towns and even cities in the Central and Western regions (29). Therefore, in the near future, China not only has to face the challenges of fast aging at the national level, but also has to balance the economic gap between the sending and receiving regions and handle a series of problems raised by internal migration and a regional aging imbalance.

Our study has several limitations. There may be some variance in the measure on the rate of net in-migration population and the natural population growth rate because of data limitations. In addition, we employed data from 2000 to 2010. It would be great if we could analyze the continuous changes when more longitudinal data sets are available. Nevertheless, the present study examined data at the city level to investigate the relationship between population aging and migration, and focused on prevalence as well as changes of aging in spatial patterns. This has promising implications for the older adults' health and income security.

5. Conclusion

The findings of our study showed the patterns of regional imbalance in aging have changed from 2000 to 2010 in China. The less developed areas have faced a faster pace of population aging within the decade than those of more advanced areas. The gap of aging level

between the other three zones and the Eastern zone was considerably narrowed down. It was the massive internal migration with obvious age selectivity and orientation preference, rather than births and deaths, that played the crucial role on the changes of regional aging patterns, which in turn brought up challenges and unfavorable consequences across regions.

Acknowledgements

This work was supported by a grant from National Natural Science Foundation of China (Grant No.71490734). We thank Professor Zuo Xuejin, Professor Gu Baochang and Professor Jin Chunlin for their valuable suggestions on the article.

References

1. Lee R. The demographic transition: Three centuries of fundamental change. *The Journal of Economic Perspectives*. 2003; 17:167-190.
2. Mu G, Chen W. Demographic transition in China: Process, characteristics and cause. *Open Time*. 2011; 1:92-101. (in Chinese)
3. Kincannon C L, He W, West LA. Demography of aging in China and the United States and the economic well-being of their older populations. *J Cross Cult Gerontol*. 2005; 20:243-255.
4. Population Reference Bureau. China's rapid aging population. *Today's Research on Aging*. 2010; 20:1-5.
5. U.S. Census Bureau. International programs center, international data base. <http://www.census.gov/ipc/www/idbnew.html>. (accessed June 14, 2017).
6. National Bureau of Statistics of China. China Statistical Yearbook 2017. <http://www.stats.gov.cn/tjsj/ndsj/2017/indexeh.htm>. (accessed June 14, 2016)
7. United Nations. World Population Prospects, the 2015 Revision. <https://esa.un.org/unpd/wpp/DataQuery/>. (accessed June 14, 2017).
8. Du P, Yang H. "Ageing before Affluence" Is a central feature of population ageing in China at the present stage. *Population Research*. 2006; 6:28-36. (in Chinese)
9. Cai F. "Ageing before Affluence" and China's sustainable economic growth. *International Economic Review*. 2012; 1:82-95.
10. Zhu H. Adult children's characteristics and intergenerational financial transfers in urban China. *Chinese Journal of Sociology*. 2016; 2:75-94.
11. Chen F, Short SE. Household context and subjective well-being among the oldest old in China. *J Fam Issues*. 2008; 29:1379-1403.
12. Cong Z, Silverstein M. Which sons live with their older parents in rural China? The role of migration and intergenerational exchanges. *Family Science*. 2010; 1:63-72.
13. Chen X, Silverstein M. Intergenerational social support and the psychological well-being of older parents in China. *Research on Aging*. 2000; 22:43-65.
14. Jones GW. Ageing in China, India and Indonesia: An Overview. In: *Contemporary Demographic Transformations in China, India and Indonesia*. Springer International Publishing, 2016; pp. 271-276.

15. Zheng Z, Yang G. Internal Migration in China: Changes and Trends. In: Contemporary Demographic Transformations in China, India and Indonesia. Springer International Publishing, 2016; pp. 223-237.
16. Zeng Yi. Urbanization and its impacts on population development in China. *Population Journal*. 1991; 2:1-6. (in Chinese)
17. Du P, Wang W. The difference of urban and rural ageing and its transition. *Population Research*. 2010; 2:3-10. (in Chinese)
18. Wang Z. The impact of rural-urban population migration on rural population aging – quantitative analysis from the view of age-rate of migration. *West Forum*. 2011; 6:27-33. (in Chinese)
19. Wu C, Wang L, Miao R. Process, prospects and countermeasures of population aging with Chinese characteristics. *Population Research*. 2004; 1:8-15. (in Chinese)
20. Zhong S, Zhao Y, Ren J. "Getting Old before Getting Rich" in China: A regional perspective. *Population Research*. 2004; 1:63-73. (in Chinese)
21. National Bureau of Statistics of China. Tabulation on the 2000 Population Census of the People's Republic of China by County. <http://www.stats.gov.cn/tjsj/ndsj/renkoupuha/2000fexian/fexian.htm>. (accessed June 14, 2016) (in Chinese)
22. Population Census Office under the State Council and Department of Population and Employment Statistics National Bureau of Statistics. Tabulation on the 2010 Population Census of the People's Republic of China by County. China Statistics Press, 2012: pp.2-407. (in Chinese)
23. National Bureau of Statistics of China. Tabulation on the 2000 Population Census of the People's Republic of China. <http://www.stats.gov.cn/tjsj/pcsj/rkpc/5rp/index1.htm>. (accessed June 14, 2017) (in Chinese)
24. National Bureau of Statistics of China. Tabulation on the 2010 Population Census of the People's Republic of China. <http://www.stats.gov.cn/tjsj/pcsj/rkpc/6rp/indexch.htm>. (accessed June 14, 2017) (in Chinese)
25. Jones GW. Ageing in China, India and Indonesia: An Overview. In: Contemporary Demographic Transformations in China, India and Indonesia. Springer International Publishing, 2016: pp.325-328.
26. Chen R. An analysis of the floating population in Shanghai based on the census data. *Manager' Journal*. 2012; 19:13-14. (in Chinese)
27. Gu B. Internal migration dominates population dynamics in China. *Asian Population Studies*. 2014; 10:1-3.
28. Zheng B. China Pension Report 2012. Economy & Management Publishing House, Beijing, China, 2012; pp.10-11. (in Chinese)
29. Zuo X. Future scenarios of population spatial distribution in China. *Social science weekly*. 2016.12.8: Version 002. (in Chinese)

(Received September 30, 2017; Revised February 11, 2018; Accepted February 28, 2018)

Intimate partner violence victimization and HIV infection among men who have sex with men in Shanghai, China

Ying Liu^{1,2,3}, Yuyan Zhang³, Zhen Ning⁴, Huang Zheng⁵, Yingying Ding^{1,2}, Meiyang Gao¹, Frank Y. Wong^{1,6,7}, Na He^{1,2,*}

¹ Department of Epidemiology, School of Public Health, and the Key Laboratory of Public Health Safety of Ministry of Education, Fudan University, Shanghai, China;

² Collaborative Innovation Center of Social Risks Governance in Health, Fudan University, Shanghai, China;

³ Putuo District Center for Disease Control and Prevention, Shanghai, China;

⁴ Shanghai Center for Disease Control and Prevention, Shanghai, China;

⁵ Shanghai PiaoXue Cultural Media Limited, Shanghai, China;

⁶ Department of Tropical Medicine, John A. Burns School of Medicine, University of Hawai'i at Mānoa, USA;

⁷ Department of Psychology, College of Social Sciences, University of Hawai'i at Mānoa, USA.

Summary

Intimate partner violence (IPV) and its association with HIV infection among men who have sex with men (MSM) in China are not understood. In this study, 732 MSM recruited from Shanghai, China between March and August 2015 were administered with a questionnaire survey and HIV blood testing. IPV victimization was measured by 25 forced-choice items capturing lifetime experience of physical, sexual, psychological, deprivation or neglect, and other forms of violence. Of them, 179 (24.3%) reported having experienced at least one type of IPV victimization. In separate multivariable analyses, sexual violence was associated with age over 35 years (AOR = 0.26, 95% CI: 0.07-1.02), ever had male-to-male commercial sex (AOR = 2.53, 95% CI: 1.19-5.39), and diagnosis of a sexually transmitted infection (STI) (AOR = 2.14, 95% CI: 0.98-4.66). Both psychological violence (AOR = 2.53, 95% CI: 1.25-5.12) and deprivation or neglect violence (AOR = 1.75, 95% CI: 1.14-2.68) were associated with ever had sex with a casual male partner(s). Having experienced at least one type of IPV victimization was significantly associated with ever had sex with a causal partner(s) (AOR = 1.72, 95% CI: 1.15-2.57) and ever had a diagnosis of a STI (AOR = 1.80, 95% CI: 1.12-2.88). HIV infection was marginally associated with having experienced any form of IPV victimization. IPV victimization is common among MSM, especially young MSM, in China, although its association with HIV infection is not conclusive in our sample. Nonetheless, our findings highlight the importance of the needs of individualized IPV interventions for certain target risk groups of MSM.

Keywords: Intimate partner violence, unprotected sex, HIV infection, MSM, China

1. Introduction

HIV has rapidly spread among men who have sex with men (MSM) in China during the past decade. The first national epidemiological survey of HIV among a total of 47,231 MSM from 61 cities from February 2008 to

September 2009 revealed an overall prevalence of 4.9% for HIV infection (1). The percentage of HIV infection in China attributable to homosexual transmission has increased from 12.2% in 2007 to 27.2% in 2015 (2). The MSM population has the highest estimate of HIV incidence among all known risk groups (e.g., injection drug users) (3).

There is a large body of epidemiological research documenting the influences of characteristics such as substance misuse and unprotected sex with multiple partners on contracting HIV among MSM in China (4-6). Meanwhile, the international literature reveals

*Address correspondence to:

Dr. Na He, Department of Epidemiology, School of Public Health, and the Key Laboratory of Public Health Safety of Ministry of Education, Fudan University, Shanghai 20032, China.

E-mail: nhe@fudan.edu.cn

that intimate partner violence (IPV), which is a subset of interpersonal violence and generally defined as occurring between spouses or other intimate partners, and encompasses multiple domains of violent behavior (e.g., physical, sexual, psychological, financial, stalking) (7), is another major determinant for contracting HIV. However, this body of IPV-HIV literature primarily focuses on women (8,9).

While IPV and its association with HIV infection and/or non-HIV sexually transmitted infections (STIs) among MSM are seldom examined (10,11), the limited literature does suggest that IPV may be a common experience among MSM with a prevalence like or even higher than that among heterosexual women, and have associations with increased odds of substance use, depressive symptoms and engagement in unprotected anal sex (9, 11-15).

To the best of our knowledge, there are only two studies examining the relationship between IPV and sexual-related risks among MSM in China. In a study on MSM in Shanghai, China, Dunkle *et al.* reported a IPV prevalence of 51% and an association of IPV with more unprotected sex and more sex linked to alcohol and other substance use (16). Another online survey in China reported that 29.8% of MSM have experienced at least 1 type of IPV, and men who have experienced IPV are more likely to have participated in group sex, to have had sex in exchange for gifts or money, and to self-report having tested positive for HIV (17). However, the association between IPV victimization and confirmed HIV infection has not been assessed in these two studies, which compromised the reliability of the association between IPV and HIV infection.

To fill this gap, we conducted a cross-sectional study to examine the prevalence of IPV victimization and its association with risk behaviors and HIV infection among MSM in Shanghai, China.

2. Materials and Methods

2.1. Study Design and Participants

Between March and August 2015, MSM clients accessing services from two voluntary HIV voluntary counseling and testing (VCT) clinics in Shanghai, China were invited to participate in a cross-sectional survey. To be eligible, potential participants must be: (1) born a biological male, (2) ages 18 years or older, (3) have had oral and/or anal sex with another man in the past 12 months, and (4) be able to give written and verbal consent. The study was approved by the Institutional Review Board of Fudan University and written informed consent was obtained from each participant.

2.2. Data Collection

Each participant was given a face-to-face interview

using a standardized questionnaire to solicit information pertaining to socio-demographic characteristics, sexual behaviors, illicit drug use, and history of STI diagnosis (e.g., syphilis, gonorrhea, chlamydia, genital herpes and genital warts). HIV testing was also performed.

2.3. Behavioral Measures

IPV from a male sexual partner (or IPV victimization) was measured by 25 items soliciting lifetime experience of physical, sexual, psychological, deprivation or neglect, and other violence from a male partner. These items were developed in accordance with the definition and typology of violence of the World Health Organization (18), the construction of items used in the WHO Multi-country Study on Women's Health and Domestic Violence against Women (19), with some having previously been validated in the study of IPV and sexual behaviors among MSM in Shanghai and in the nationwide online survey of sexual health among MSM in China (16,20). These 25 forced-choice items cover five domains: physical violence (7 items), sexual violence (3 items), psychological violence (5 items), deprivation or neglect violence (7 items), and other violence such as having ever been threatened to reveal their sexuality (3 items).

2.4. HIV Testing

Serum samples were screened for anti-HIV IgG antibody by an enzymatic link immunosorbent assay (ELISA) (Wantai Bio-Pharm, Beijing, China). All samples that screened positive for HIV were subject to another antibody testing with another HIV diagnostic kit using ELISA tests. The positive samples were then confirmed by a Western Blot Assay (HIV BLOT 2.2; Genelabs Diagnostics Pte Ltd., Singapore). All laboratory tests were performed according to the manufacturers' instructions.

2.5. Data Analysis

The distributions of socio-demographic characteristics, sexual behaviors, illicit drug use, diagnosis of any STIs, and IPV victimization were tabulated as percentages. HIV prevalence was compared using chi-square test or Fisher's exact test when appropriate. Univariable and multivariable logistic regression analyses were performed separately to explore sociodemographic and sexual behavioral correlates of various types of IPV victimization. To examine the correlation between IPV victimization and HIV infection, multiple logistic regression model was employed with adjustment for demographic variables. All statistical analyses were performed using SPSS software (SPSS Inc., Chicago, Illinois, USA).

3. Results

3.1. Socio-demographic Characteristics

A total of 760 MSM were approached and identified to be eligible. Of them, 732 (96.3%) participated in the study. Among the participants, 83.9% were younger than 35 years old; 80.3% were never married; 59.7% were migrants (*i.e.*, not born in Shanghai); 78.3% had college education; and 73.6% self-identified as gay. Approximately 9.8% of the participants had ever used illicit drugs (Table 1).

3.2. Homosexual Behaviors

Fifteen-percent of the participants had ever engaged in male-to-male commercial sex. Participants reported a high prevalence (61.9%) of lifetime experience of sex with a casual male partner(s). Nearly two-thirds (66.9%) of the participants reported having had two or more male sex partners in the past year. About 14.1% reported having ever been diagnosed with a STI.

3.3. Prevalence of IPV Victimization

Among the participants, the overall prevalence of any type IPV victimization suffering from a male sexual partner(s) was 24.3%. Specifically, the prevalence was 6.6% for any form of physical violence; 5.5% for any form of sexual violence; 8.2% for any form of psychological violence; 20.1% for any form of deprivation or neglect violence, and 2.3% for any form of other violence (Table 2).

3.4. Correlates of IPV Victimization

As shown in Table 3, multivariable analyses reveal that MSM (1) who ever had commercial sex with a male sex worker(s) (AOR = 2.53, 95%CI: 1.19 to 5.39) and a diagnosis of any type of STI (AOR = 2.14, 95% CI: 0.98-4.66) reported more sexual violence; (2) who aged 35 years and older reported less sexual violence (AOR = 0.26, 95% CI: 0.07 to 1.02); (3) who ever had sex with a casual male partner(s) in lifetime reported more psychological violence (OR = 2.53, 95% CI: 1.25 to 5.12) and deprivation or neglect violence (AOR = 1.75, 95% CI: 1.14 to 2.68). Other forms of violence were positively associated with a diagnosis of any STI (AOR = 3.28, 95% CI: 1.05 to 10.27), but negatively associated with having college or advanced education (AOR = 0.23, 95% CI: 0.06 to 0.93) and MSM who self-identified as gay (AOR = 0.23, 95%CI: 0.08 to 0.71). Having experience of any type of IPV victimization was positively associated with having sex with a casual male partner(s) (AOR = 1.72, 95% CI: 1.15 to 2.57) and diagnosis of any STI (AOR = 1.80, 95% CI: 1.12 to 2.88). Lifetime illicit drug was marginally

associated with physical violence.

3.5. HIV Infection and Its Association with IPV Victimization

Eighty-five participants (11.6%) were tested positive for HIV. The HIV prevalence was significantly higher among participants with lower education, HIV-risk taking behaviors including commercial sex, casual sex, multiple partnership and illicit drug use, as well as diagnosis of a STI (Table 1). As shown in Table 2, the HIV prevalence was 11.4% for those without any IPV, 20.0% for those with one type of IPV and 9.4% for those with two or more types of IPV. No separate and significant association was identified between HIV infection and any of the five violence domains using bivariate chi-square test (Table 2). However, there was a marginally significant association between HIV infection and IPV after adjusted for age, education, marital status and sexual identity according to a multivariable logistic regression analysis (Table 4).

4. Discussion

In this study we examined IPV among MSM in China, and its association to HIV infection. Our finding reveals a relatively high prevalence of IPV victimization among MSM attending VCT clinics in Shanghai, China. It is similar to the rate reported by the nationwide online survey among MSM in China (17), but lower than the rate among MSM in Shanghai reported by Dunkle *et al* (16). The possible explanation for the discrepant results between our study and the Dunkle's is that about a half of participants in Dunkle's study were "money boys" (*i.e.*, man who sells sex to man), who may be more vulnerable to IPV than general MSM.

The difference in estimating IPV prevalence may also reflect differences in sampling methodology used in the three studies. We used a clinic-based methodology to recruit MSM, whereas Dunkle *et al.* used Respondent-driven sampling methodology which has been shown for attracting higher risk populations such as "money boys." Indeed, Ibragimov *et al.* argue that type of recruitment method used is likely to reach a certain sub-type of MSM with a specific risk profile (21); therefore, these researchers recommend multiple methods are needed to obtain a representative sample of MSM.

Consistent with the known literature(11,22-26), we also found that MSM engaging in high-risk sexual behaviors such as commercial sex and casual sex, or had a history of STIs are more likely to report IPV victimization. Notably, our findings reveal that young MSM reported more sexual violence than older MSM. One possible explanation for this phenomenon might be due to sexual mixing with older male partners among young MSM in Shanghai, partly for emotional,

Table 1. Demographic and sexual behavioral characteristics, lifetime intimate partner violence and HIV infection among MSM in Shanghai, China

Characteristics	No. (%) ^a	Physical violence No. (%) ^a	Sexual violence No. (%) ^a	Psychological violence No. (%) ^a	Deprivation/neglect violence No. (%) ^a	Other violence No. (%) ^a	Any violence No. (%) ^a	HIV infection No. (%) ^b	P Value ^c
Age, years									0.098
18-25	266 (36.3)	13 (4.9)	15 (5.6)	16 (6.0)	57 (21.4)	6 (2.3)	67 (25.2)	22 (8.3)	
26-35	348 (47.6)	29 (8.3)	21 (6.0)	36 (10.3)	73 (21.0)	9 (2.6)	88 (25.3)	48 (13.8)	
36-74	118 (16.1)	6 (5.1)	4 (3.4)	8 (6.8)	17 (14.4)	2 (1.7)	23 (19.5)	15 (12.7)	
Marital status									0.068
Never married/ Ever married	588 (80.3) 144 (19.7)	37 (6.3) 11 (7.6)	31 (5.3) 9 (6.2)	48 (8.2) 12 (8.3)	122 (20.7) 25 (17.4)	13 (2.2) 4 (2.8)	147 (25.0) 31 (21.5)	62 (10.5) 23 (16.0)	
Education									< 0.001
Middle school High school College or above	49 (6.7) 110 (15.0) 573 (78.3)	5 (10.2) 10 (9.1) 33 (5.8)	4 (8.2) 11 (10.0) 25 (4.4)	7 (14.3) 9 (8.2) 44 (7.7)	11 (22.4) 26 (23.6) 110 (19.2)	4 (8.2) 5 (4.5) 8 (1.4)	14 (28.6) 32 (29.1) 132 (23.0)	14 (28.6) 14 (12.7) 57 (9.9)	
Self-identified sexual identity									0.528
Gay Non-gay	539 (73.6) 193 (26.4)	15 (7.8) 33 (6.1)	8 (4.1) 32 (5.9)	16 (8.3) 44 (8.2)	37 (19.2) 110 (20.4)	10 (5.2) 7 (1.3)	41 (21.2) 137 (25.4)	65 (12.1) 20 (10.4)	
Ever had commercial sex with male sexual workers									0.001
No Yes	622 (85.0) 110 (15.0)	40 (6.4) 8 (7.3)	27 (4.3) 13 (11.8)	53 (8.5) 7 (6.4)	124 (19.9) 23 (20.9)	12 (1.9) 5 (4.5)	144 (23.2) 34 (30.9)	62 (10.0) 23 (20.9)	
Ever had sex with casual male partners									0.046
No Yes	279 (38.1) 453 (61.9)	12 (4.3) 36 (7.9)	9 (3.2) 31 (6.8)	11 (3.9) 49 (10.8)	40 (14.3) 107 (23.6)	4 (1.4) 13 (2.9)	48 (17.2) 130 (28.7)	24 (8.6) 61 (13.5)	
No. of male sex partners in the past 12 months									0.025
1 2-4 ≥ 5	242 (33.1) 380 (51.9) 110 (15.0)	11 (4.5) 29 (7.6) 8 (7.3)	10 (4.1) 21 (5.5) 9 (8.2)	14 (5.8) 33 (8.7) 13 (11.8)	39 (16.1) 83 (21.8) 25 (22.7)	3 (1.2) 9 (2.4) 5 (4.5)	45 (18.6) 97 (25.5) 36 (32.7)	27 (11.2) 37 (9.7) 21 (19.1)	
Ever used illicit drugs									< 0.001
No Yes	660 (90.2) 72 (9.8)	37 (5.6) 11 (15.3)	32 (4.8) 8 (11.1)	52 (7.9) 8 (11.1)	129 (19.5) 18 (25.0)	13 (2.0) 4 (5.6)	155 (23.5) 23 (31.9)	61 (9.2) 24 (33.3)	
Ever been diagnosed with any STI									0.008
No Yes	629 (85.9) 103 (14.1)	36 (5.7) 12 (11.7)	28 (4.5) 12 (11.7)	50 (7.9) 10 (9.7)	119 (18.9) 28 (27.2)	10 (1.6) 7 (6.8)	140 (22.3) 38 (36.9)	65 (10.3) 20 (19.4)	

STI, sexually transmitted infection. a: proportion; b: prevalence; c: P value for HIV infection.

Table 2. Lifetime experience of various type of IPV victimization from male sexual partners among MSM in Shanghai

Type of IPV victimization	No. (%) ^a	HIV infection N No. (%) ^b
Domain 1: Physical violence		
Ever been slapped	24 (3.3)	
Ever been thrown with something	26 (3.6)	
Ever been pushed or shoved	25 (3.4)	
Ever been hit with fist or something else could hurt	18 (2.5)	
Ever been kicked or dragged	12 (1.6)	
Ever been choked or burnt on purpose	5 (0.7)	
Ever been threatened to use or actually used a gun, knife or other weapon against you	3 (0.4)	
Having any physical violence		<i>P</i> = 0.789
No	684 (93.4)	80 (11.7)
Yes	48 (6.6)	5 (10.4)
Domain 2: Sexual violence		
Ever been physically forced to have sexual intercourse when you did not want to	35 (4.8)	
Ever had sexual intercourse when you did not want to because of what he might do	8 (1.1)	
Ever been forced to do something sexual that degrading or humiliating	8 (1.1)	
Having any sexual violence		<i>P</i> = 0.211 ^c
No	692 (94.5)	78 (11.3)
Yes	40 (5.5)	7 (17.5)
Domain 3: Psychological violence		
Ever been insulted or made feel bad	38 (5.2)	
Ever been belittled or humiliated in front of other people	23 (3.1)	
Ever been scared or intimidated on purpose	27 (3.7)	
Ever been verbally threatened to harm you physically or emotionally	17 (2.3)	
Ever been verbally threatened to physically harm someone you care for	12 (1.6)	
Having any psychological violence		<i>P</i> = 0.212
No	672 (91.8)	81 (12.1)
Yes	60 (8.2)	4 (6.7)
Domain 4: Deprivation or neglect violence		
Ever been kept from seeing friends	31 (4.2)	
Ever been restricted to contact with your family	9 (1.2)	
Ever insisted on knowing where you are at all times	82 (11.2)	
Ever been ignored and treated indifferently	62 (8.5)	
Ever been angry with you if you spoke with another man	75 (10.2)	
Been often suspicious of your unfaithful	73 (10.0)	
Ever been expected to ask permission before seeking health care	25 (3.4)	
Having any deprivation or neglect violence		<i>P</i> = 0.241
No	585 (79.9)	72 (12.3)
Yes	147 (20.1)	13 (8.8)
Domain 5: Other violence		
Ever been threatened to reveal your sexuality	2 (0.3)	
Your property has ever been damaged or destroyed	12 (1.6)	
Ever been threatened to stop helping you with money or with housing	9 (1.2)	
Having any other violence		<i>P</i> = 0.709 ^c
No	715 (97.7)	84 (11.7)
Yes	17 (2.3)	1 (5.9)
No. of above-mentioned types of IPV		
0	554 (75.7)	63 (11.4)
1	50 (6.8)	10 (20.0)
≥ 2	128 (17.5)	12 (9.4)

^a: proportion; ^b: prevalence; ^c: Fisher's exact test.

financial and instrumental support (27). In such relationship, young MSM may be at a disadvantage position and thereby are more likely to be victims of sexual violence. Sex with older male partners may also increase their risk of HIV infection (27).

The relationship between IPV and HIV infection among MSM population has been reported in some but not all studies (10,28-30). In our sample, HIV infection is marginally associated with having experienced any form of IPV victimization. Nonetheless, our descriptive data indicate that higher prevalence of HIV infection is only observed among those with sexual violence compared with those without, suggesting that the link between IPV and HIV infection is mostly related to sexual violence. The discrepancies between studies might be due to the confluence of various types of IPV and its impact on HIV

risks. For example, an initial act of physical violence might be likely to escalate sexual violence among intimate partners. Further research is needed to elucidate the dynamics of IPV "penetration" on victimization.

The current study had several limitations. First, all participants were recruited from VCT clinics in Shanghai, limiting the ability to generalize the findings to other MSM populations in China. Second, due to the cross-sectional nature, the temporal order between IPV victimization and HIV infection could not be determined. Third, our data were self-reported, some sensitive topics such as sexual behaviors and IPV victimization experiences might be underreported by participants due to the social desirability; and these were lifetime not recent experiences and behaviors, thereby recall bias could not be neglected particularly for older MSM

Table 3. Separate multivariable logistic regression models exploring correlates of various type of IPV victimization from male sexual partners among MSM in Shanghai

Characteristics	Adjusted OR (95% CI)					
	Model 1: Physical violence	Model 2: Sexual violence	Model 3: Psychological violence	Model 4: Deprivation/neglect violence	Model 5: Other violence	Model 6: Any violence
Age, years						
18-25	1.00	1.00	1.00	1.00	1.00	1.00
26-35	1.56 (0.76-3.18)	0.79 (0.37-1.69)	1.70 (0.90-3.23)	0.91 (0.60-1.37)	0.96 (0.30-3.12)	0.90 (0.60-1.33)
36-74	0.78 (0.23-2.58)	0.26 (0.07-1.02)	0.98 (0.34-2.84)	0.52 (0.25-1.08)	0.56 (0.07-4.31)	0.59 (0.30-1.15)
Education						
Middle school	1.00	1.00	1.00	1.00	1.00	1.00
High school	0.94 (0.29-3.04)	1.32 (0.38-4.63)	0.63 (0.21-1.86)	1.14 (0.50-2.61)	0.57 (0.13-2.45)	1.11 (0.51-2.41)
College or above	0.55 (0.19-1.60)	0.55 (0.17-1.84)	0.49 (0.19-1.26)	0.79 (0.37-1.68)	0.23 (0.06-0.93)*	0.79 (0.39-1.61)
Ever married	1.12 (0.44-2.86)	1.86 (0.65-5.30)	0.97 (0.40-2.33)	0.98 (0.52-1.84)	0.54 (0.11-2.54)	0.95 (0.52-1.72)
Self-identified as gay	0.84 (0.41-1.72)	2.25 (0.90-5.64)	1.07 (0.55-2.08)	1.09 (0.68-1.72)	0.23 (0.08-0.71)**	1.31 (0.84-2.04)
Ever had commercial sex with male sex workers	0.79 (0.34-1.82)	2.53 (1.19-5.39)*	0.60 (0.26-1.40)	0.99 (0.59-1.67)	1.50 (0.45-4.98)	1.41 (0.88-2.26)
Ever had sex with casual male partner(s)	1.44 (0.70-2.96)	1.84 (0.81-4.18)	2.53 (1.25-5.12)**	1.75 (1.14-2.68)*	1.05 (0.30-3.65)	1.72 (1.15-2.57)**
No. of male sex partners in the past year						
1	1.00	1.00	1.00	1.00	1.00	1.00
2-4	1.59 (0.75-3.37)	1.16 (0.51-2.63)	1.31 (0.68-2.56)	1.30 (0.84-2.01)	2.25 (0.55-9.24)	1.3 (0.86-1.98)
≥ 5	1.03 (0.37-2.82)	1.19 (0.43-3.28)	1.43 (0.61-3.33)	1.14 (0.63-2.09)	2.57 (0.52-12.78)	1.53 (0.88-2.65)
Ever used illicit drugs	2.20 (0.99-4.88)	1.17 (0.46-2.99)	1.05 (0.45-2.44)	0.99 (0.54-1.81)	1.43 (0.35-5.86)	0.95 (0.53-1.69)
Ever been diagnosed with a STI	1.63 (0.77-3.43)	2.14 (0.98-4.66)	0.99 (0.47-2.13)	1.47 (0.89-2.44)	3.28 (1.05-10.27)*	1.80 (1.12-2.88)*

CI, confidence interval; STI, sexually transmitted infection; OR, odds ratio. * $P < 0.05$, ** $P < 0.01$.

Table 4. Multivariable logistic regression analysis of correlates with HIV infection among MSM in Shanghai

Items	Adjusted OR (95%CI)	P Value
Age, years		
18-25	1.0	
26-35	1.81 (1.04-3.16)	0.037
36-74	1.02 (0.43-2.41)	0.963
Education		
Middle school	1.0	
High school	0.39 (0.17-0.92)	0.030
College or above	0.25 (0.12-0.53)	< 0.001
Ever married	1.58 (0.77-3.24)	0.214
Self-identified as gay	1.66 (0.90-3.06)	0.105
No. of types of IPV victimization		
None	1.0	
1	2.01 (0.94-4.32)	0.072
≥ 2	0.74 (0.38-1.43)	0.363

CI, confidence interval; OR, odds ratio.

participants.

In conclusion, our findings demonstrate a high prevalence of IPV victimization among MSM particularly young MSM population in China. Our data also show that various types of IPV victimization are linked to different socio-demographic and sexual behavioral characteristics, suggesting different mechanisms of various types of IPV. These, taken together, underscore the need of individualized IPV interventions for certain target risk groups of MSM. Further research is needed to clarify the complex link between IPV, drug use, types of sexual relationship, and HIV/STI infection.

Acknowledgements

This study was supported by the Shanghai Municipal Health and Family Planning Commission (grant no. GWTD2015S05, 15GWZK0101, 201540066) and the Natural Science Foundation of China (grant no. 81361120385).

References

1. Wu Z, Xu J, Liu E, *et al.* HIV and syphilis prevalence among men who have sex with men: A cross-sectional survey of 61 cities in China. *Clin Infect Dis*. 2013; 57:298-309.
2. National Health and Family Planning Commission of The People's Republic of China. *China AIDS Response Progress Report*. Beijing, China, 2015.
3. Cui Y, Guo W, Li D, Wang L, Shi CX, Brookmeyer R, Detels R, Ge L, Ding Z, Wu Z. Estimating HIV incidence among key affected populations in China from serial cross-sectional surveys in 2010-2014. *J Int AIDS Soc* 2016; 19:20609.
4. Nehl EJ, He N, Lin L, Zheng T, Harnisch JA, Ding Y, Wong FY. Drug use and sexual behaviors among MSM in China. *Subst Use Misuse*. 2015; 50:123-136.
5. Yu F, Nehl EJ, Zheng T, He N, Berg CJ, Lemieux AF, Lin L, Tran A, Sullivan PS, Wong FY. A syndemic including cigarette smoking and sexual risk behaviors among a sample of MSM in Shanghai, China. *Drug Alcohol*

- Depend. 2013; 132:265-270.
6. Huang ZJ, He N, Nehl EJ, Zheng T, Smith BD, Zhang J, McNabb S, Wong FY. Social network and other correlates of HIV testing: Findings from male sex workers and other MSM in Shanghai, China. *AIDS Behav.* 2012; 16:858-871.
 7. Finneran C, Stephenson R. Intimate Partner Violence Among Men Who Have Sex With Men: A Systematic Review. *Trauma Violence Abuse.* 2013; 14:168-185.
 8. Jewkes RK, Dunkle K, Nduna M, Shai N. Intimate partner violence, relationship power inequity, and incidence of HIV infection in young women in South Africa: A cohort study. *Lancet.* 2010; 376:41-48.
 9. Li Y, Marshall CM, Rees HC, Nunez A, Ezeanolue EE, Ehiri JE. Intimate partner violence and HIV infection among women: A systematic review and meta-analysis. *J Int AIDS Soc.* 2014; 17:18845.
 10. Dunkle KL, Jewkes RK, Murdock DW, Sikweyiya Y, Morrell R. Prevalence of Consensual Male-Male Sex and Sexual Violence, and Associations with HIV in South Africa: A Population-Based Cross-Sectional Study. *PLoS Med.* 2013; 10:e1001472.
 11. Buller AM, Devries KM, Howard LM, Bacchus LJ. Associations between intimate partner violence and health among men who have sex with men: A systematic review and meta-analysis. *PLoS Med.* 2014; 11:e1001609.
 12. Duncan DT, Goedel WC, Stults CB, Brady WJ, Brooks FA, Blakely JS, Hagen D. A study of intimate partner violence, substance abuse, and sexual risk behaviors among gay, bisexual, and other men who have sex with men in a sample of geosocial-networking smartphone application users. *Am J Mens Health.* 2018; 12:292-301.
 13. Williams JK, Wilton L, Magnus M, Wang L, Wang J, Dyer TP, Koblin BA, Hucks-Ortiz C, Fields SD, Shoptaw S, Stephenson R, O Cleirigh C, Cummings V. Relation of childhood sexual Abuse, intimate partner violence, and depression to risk factors for HIV among black men who have sex With men in 6 US Cities. *Amer J Public Health.* 2015; 105:2473-2481.
 14. Rhodes SD, McCoy TP, Wilkin AM, Wolfson M. Behavioral risk disparities in a random sample of self-identifying gay and non-gay male university Students. *J Homosexuality.* 2009; 56:1083-1100.
 15. Stockman JK, Lucea MB, Campbell JC. Forced sexual initiation, sexual intimate partner violence and HIV risk in women: A global review of the literature. *AIDS Behav.* 2013; 17:832-847.
 16. Dunkle KL, Wong FY, Nehl EJ, Lin L, He N, Huang J, Zheng T. Male-on-male intimate partner violence and sexual risk behaviors among money boys and other men who Have sex with men in Shanghai, China. *Sex Transm Dis.* 2013; 40:362-365.
 17. Davis A, Best J, Wei C, Luo J, Van Der Pol B, Meyerson B, Dodge B, Aalsma M, Tucker J. Intimate partner violence and correlates with risk behaviors and HIV/STI diagnoses among men who have sex with men and men who have sex with men and women in China. *Sex Transm Dis.* 2015; 42:387-392.
 18. World Health Organization (Internet). Definition and typology of violence. <http://www.who.int/violenceprevention/approach/definition/en/> (accessed Dec 20, 2016).
 19. Garcia-Moreno C, Jansen HA, Ellsberg M, Heise L, Watts CH. Prevalence of intimate partner violence: Findings from the WHO multi-country study on women's health and domestic violence. *Lancet.* 2006; 368:1260-1269.
 20. Zhang Y, Best J, Tang W, Tso LS, Liu F, Huang S, Zheng H, Yang B, Wei C, Tucker JD. Transgender sexual health in China: A cross-sectional online survey in China. *Sex Transm Infect.* 2016; pii:sextrans-2015-052350.
 21. Ibragimov U, Harnisch JA, Nehl EJ, He N, Zheng T, Ding Y, Wong FY. Estimating self-reported sex practices, drug use, depression, and intimate partner violence among MSM in China: A comparison of three recruitment methods. *AIDS Care.* 2017; 29:125-131.
 22. Kramer SC, Schmidt AJ, Berg RC, Furegato M, Hospers H, Folch C, Marcus U. Factors associated with unprotected anal sex with multiple non-steady partners in the past 12 months: Results from the European Men-Who-Have-Sex-With-Men Internet Survey (EMIS 2010). *BMC Public Health.* 2015; 16:47.
 23. Koblin BA, Torian L, Xu G, Guilin V, Makki H, Mackellar D, Valleroy L. Violence and HIV-related risk among young men who have sex with men. *AIDS Care.* 2006; 18:961-967.
 24. Houston E, McKirnan DJ. Intimate partner abuse among gay and bisexual men: Risk correlates and health outcomes. *J Urban Health.* 2007; 84:681-690.
 25. Decker MR, Seage GR, Hemenway D, Gupta J, Raj A, Silverman JG. Intimate partner violence perpetration, standard and gendered STI/HIV risk behaviour, and STI/HIV diagnosis among a clinic-based sample of men. *Sex Transm Infect.* 2009; 85:555-560.
 26. Raj A, Kidd JD, Cheng DM, Coleman S, Bridden C, Blokhina EA, Krupitsky E, Samet JH. Associations between partner violence perpetration and history of STI among HIV-infected substance using men in Russia. *AIDS Care.* 2013; 25:646-651.
 27. Ding Y, Zhou Y, Liu C, Liu X, He N. Sex with older partners, condomless anal sex and unrecognized HIV infection among Chinese men who have sex with men. *AIDS Care.* 2018; 30:305-311.
 28. Stults CB, Javdani S, Greenbaum CA, Kapadia F, Halkitis PN. Intimate partner violence and substance use risk among young men who have sex with men: The P18 cohort study. *Drug Alcohol Depend.* 2015; 154:54-62.
 29. Wirtz AL, Jumbe V, Trapence G, Kamba D, Umar E, Ketende S, Berry M, Strömdahl S, Beyrer C, Baral SD. HIV among men who have sex with men in Malawi: Elucidating HIV prevalence and correlates of infection to inform HIV prevention. *J Int AIDS Soc.* 2013; 16(Suppl 3):18742.
 30. Relf MV, Huang B, Campbell J, Catania J. Gay Identity, Interpersonal violence, and HIV risk behaviors: An empirical test of theoretical relationships among a probability-based sample of urban men who Have Sex with Men. *J Assoc Nurses AIDS Care.* 2004; 15:14-26.

(Received March 10, 2018; Revised April 22, 2018; Accepted April 23, 2018)

Astragalus polysaccharide protects diabetic cardiomyopathy by activating NRG1/ErbB pathway

Xiao Chang¹, Kang Lu², Ling Wang¹, Min Lv¹, Wenjun Fu^{3,*}

¹Department of intensive care unit, Shenzhen Traditional Chinese Medicine Hospital, Shenzhen, Guangdong, China;

²School of Basic Medical Science, Guangzhou University of Chinese Medicine, Guangzhou, Guangdong, China;

³South China Research Center for Acupuncture and Moxibustion, School of Basic Medical Science, Guangzhou University of Chinese Medicine, Guangzhou, Guangdong, China.

Summary

Diabetic cardiomyopathy (DCM) is one of the main cardiac complications among diabetic patients. According to previous studies, the pathogenesis of DCM is associated with oxidative stress, apoptosis and proliferation of local cardiac cells. It showed, NRG1 can improve the function of mitochondria, and thereby, increasing proliferation and decreasing apoptosis of cardiac muscle cell *via* ErbB/AKT signaling, also, exert antioxidative function. Besides, NRG1/ErbB pathway was impaired in the DCM model which suggested this signaling played key role in DCM. *Astragalus polysaccharide* (APS), one of the active components of *Astragalus mongholicus*, showed striking antioxidative effect. Here, in this study, our data showed that APS can promote proliferation and decrease apoptosis in AGE-induced DCM cell model, besides, APS can decrease intracellular ROS level, increase activity of SOD, GSH-Px and lower level of MDA and NO in DCM cell model, indicating APS exerted antioxidative function in DCM model cells. Besides, western blot results revealed APS induced NRG1 expressing and the phosphorylation level of ErbB2/4. In addition, the elevated NRG1 promoted AKT and PI3k phosphorylation which indicated APS may exert its function by NRG1/ErbB and the downstream AKT/PI3K signaling. Canertinib is ErbB inhibitor. The effect of APS on proliferation, apoptosis, antioxidation and NRG1/ErbB pathway was partly abolished after the cells were co-treated with APS and canertinib. Taken together, these results suggested APS may display its protective function in DCM cells by activating NRG1/ErbB signaling pathway. And our study increased potential for prevention and therapy to DCM.

Keywords: *Astragalus polysaccharide* (APS), Diabetic cardiomyopathy, Antioxidation, NRG1/ErbB

1. Introduction

Diabetes mellitus (DM) is the most common metabolic disorders worldwide. In 2010, the diabetic patients were 285 million and, it will increase to 439 million by 2030 (1,2). Diabetic patients were characterized with persistent hyperglycemia which may lead to damage to various organs such as the heart (1,2). Therefore, DM may cause

cardiovascular complications including coronary heart disease, hypertension and diabetic cardiomyopathy (DCM) which are responsible for 80% of mortality and morbidity for diabetic patients (3-6). The characterized pathology of DCM is distinct from hypertension and coronary artery diseases (3-6). According to the previous studies, the main reason for pathological change of DCM is microangiopathy, which can cause damage to cardiac structure and function, such as apoptosis of the cardiac cells (3-6). But, the pathogenesis of DCM has not been clearly understood.

It's showed that various biology processes are associated with DCM, such as cardiomyocyte apoptosis, oxidative stress which is believed to be the key reason induced DM to DCM (7-10). There is a balance between reactive nitrogen species and reactive oxygen species

Released online in J-STAGE as advance publication April 2, 2018.

*Address correspondence to:

Dr. Wenjun Fu, South China Research Center for Acupuncture and Moxibustion, School of Basic Medical Science, Guangzhou University of Chinese Medicine, No. 232, East Waihuan Road, Panyu District, Guangzhou, Guangdong 510006, China.
E-mail: fuqingzhu2006@163.com; fuwenjun201511@163.com

production in normal cells, while, when the cells were progressed to DCM, the balance was disrupted and the accumulated production can't be cleaned in time which finally caused damage. Reactive oxygen species (ROS) are reactive chemical species and excess accumulated ROS can induce oxidative stress (11-16). Hence, inhibition of the oxidative stress and improvement of the antioxidative function are believed to be one of the important therapeutic strategies for DCM patients.

The studies showed that antioxidative natural products displayed potential therapeutic effect for DCM (4,17-21). Among these natural products, *Astragalus polysaccharides* (APS) is one of the main active extract from *Astragalus membranaceus* (20,22-27). Previous studies showed that, for diabetic cardiomyopathy in hamsters, APS can induce myocardial collagen deposition and improve cardiac function by activating ERK1/2 signal pathway (23,28). Besides, APS can improve cardiac glucose metabolism dysfunction by inducing expression of GLUT-4 and inhibiting PPAR expressing (20,24,29).

Here, in this study, we showed that, APS can promote proliferation and inhibit apoptosis in AGE-induced H9C2 DCM model cells. Besides, APS exerts antioxidative function. Further studies showed, APS exerts its promoting proliferation, suppressing apoptosis and antioxidative function by NRG1/ErbB and its downstream AKT/PI3K pathway in DCM model cells. In summary, our study proved APS exerts protective function in DCM model cells by NRG1/ErbB signal pathway which suggesting APS have great promising for DCM therapy.

2. Materials and Methods

2.1. Cell culture and treatment

H9C2 cells (purchased from Shanghai Cellular Research Institute, Shanghai, China) were cultured with Dulbecco's modified essential medium (DMEM, Hyclone, USA) supplemented with 10% fetal bovine serum (FBS, Gibco, USA) and 1% antibiotic-antimycotic (Gibco, USA) and were cultured at 37°C in 5% CO₂ incubator. To establish DCM model, H9C2 cells were treated with 200 mg/L AGE-BSA and 30 mm/L glucose for 24 h, besides, the cells treated with 200 mg/L BSA and 5 mm/L glucose were used as normalized control group. For the experimentation, the model cells were treated with indicated APS (Tianjin Cinorch Pharmaceutical Company, China) or APS combined with NRG1/ErbB1 inhibitors Canertinib (Sigma, USA).

2.2. CCK8 assay

Cell proliferation ability was tested by CCK8 assay (Beyotime Biotechnology, China). The normalized control group or diabetes mellitus model group H9C2

cells in 96-well plates were treated with the indicated concentration APS for indicated time or APS combined with Canertinib, then the cells were incubated with CCK-8 solution at 37°C for 2 hours. Subsequently, the optical density (OD) at 450 nm wavelength was tested by Microplate reader (Biotek, USA).

2.3. ROS assay

The intracellular ROS production was detected by ROS assay Kit (Beyotime, China) and the measurement process was carried according to the protocol. Following the indicated treatment, the H9C2 cells in 96-well were washed 2 times with PBS and then incubated with 10 μmol/L 2, 7- dichlorodi-hydrofluorescein diacetate (DCFH-DA) for 20 minutes at 37°C. Followingly, the cells were washed three times with PBS and then the ROS-sensitive signal was examined with Microplate reader at excitation wavelength of 488 nm and emission wavelength of 525 nm.

2.4. ELISA assay

The activity of superoxide dismutase (SOD) and Glutathione peroxidase (GSH-Px) and level of malondialdehyde (MDA) and NO₁ in cell supernatant were tested by the respective ELISA kit (Jiancheng Bioengineering Institute, China) according to the manufacturer's protocol. The H9C2 cells were treated as indicated and the cell supernatant was collected. After the treatment process, the absorbance was measured by spectrophotometer.

2.5. Apoptosis analysis

Cell apoptosis was analyzed by flow cytometry. After indicated treatment, the cells were collected, washed with PBS solution and then resuspended in binding buffer. Then, then cells were stained with PI and Annexin V (Invitrogen, USA) for 15 minutes in dark at room temperature. Subsequently, the doubled stained cells were analyzed by flow cytometer (BD, USA).

2.6. Western blot

The protein samples were harvested by loading buffer. The proteins were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and then transferred to polyvinylidene difluoride (PVDF) membranes (Millipore, Germany). After been blocked with 5% non-fat milk TBST solution for 2 hours, the membranes were incubated overnight at 4°C with primary antibodies against NRG1 (Abcam, 1:1,000), p-ErbB2 (Abcam, 1:800), ErbB2 (Abcam, 1:1,000), p-ErbB4 (Abcam, 1:2,000), ErbB4 (Abcam,1:500), p-PI3K (Abcam,1:2,000), PI3K (Abcam, 1:1,000), p-AKT (Abcam,1:500), AKT (Abcam,1:500) and

GAPDH (Abcam, 1:10,000). Following with secondary antibody (BOSTER, 1:20,000) for 1 hour, the membranes were visualized by ECL kit (Thermo, USA) and the image was scanned and collected by ScanMaker 1000XL Plus instrument.

2.7. Statistical analysis

All data were presented as mean \pm SD. Every experiment was replicated at least three times. Student's *t*-test and one-way ANOVA were used to perform statistical analysis. *P* value < 0.05 was considered to have statistical difference.

3. Results

3.1. APS promotes proliferation and inhibits oxidative stress in DCM model H9C2 cells

To identify the function of APS in DCM, firstly, we tried to get DCM model cells. H9C2 cells were incubated with 200 mg/L AGE-BSA and 30 mm/L glucose for 24 hours to obtain DCM model cells and the treated with 200 mg/L BSA and 5 mm/L glucose cells were used as normal control (NC). The DCM model cells were cultured with various concentrations of APS (0.1, 1.0, 10, 100 $\mu\text{g}/\text{mL}$) for indicated time (0, 12, 24, 48 and 72 hours). After the treatment, cell viabilities were tested by CCK-8 assay. The data showed, the APS can increase the cell viabilities dose-dependently and time-dependently (Figure 1A). Besides, compared with NC group cells, the intracellular ROS level was higher in DCM model cells (Figure 1B) which is consistency with the previous studies, while, APS decreased intracellular ROS level in DCM model cells dose-dependently and time-dependently (Figure 1B). DCM is closely associated with oxidative stress.

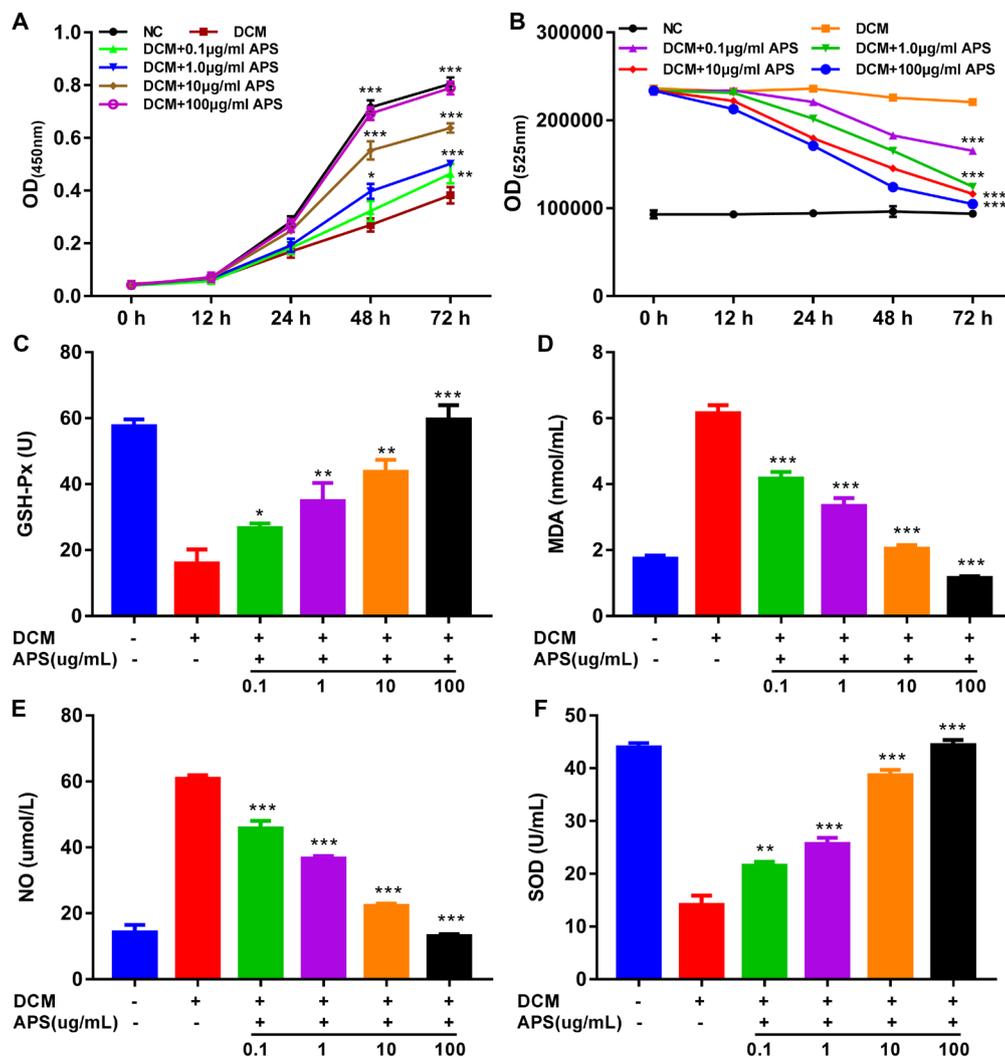


Figure 1. APS promotes proliferation and inhibits oxidative stress in DCM model H9C2 cells. (A and B) DCM model H9C2 cells were treated with indicated concentrations of APS for various time. CCK-8 assay was used to test the cell viability and intracellular ROS level was tested by DCF. NC (normal control), DCM (DCM model cells). (C, D, E, and F) The DCM model H9C2 cells were incubated with indicated concentrations APS for 48 hours. The activity of GSH-Px (C) and SOD (F) and the level of MDA (D) and NO (E) were examined by ELISA assay. Data was presented as Mean \pm SD, **p* < 0.05 , ***p* < 0.01 , and ****p* < 0.001 vs. DCM group cells.

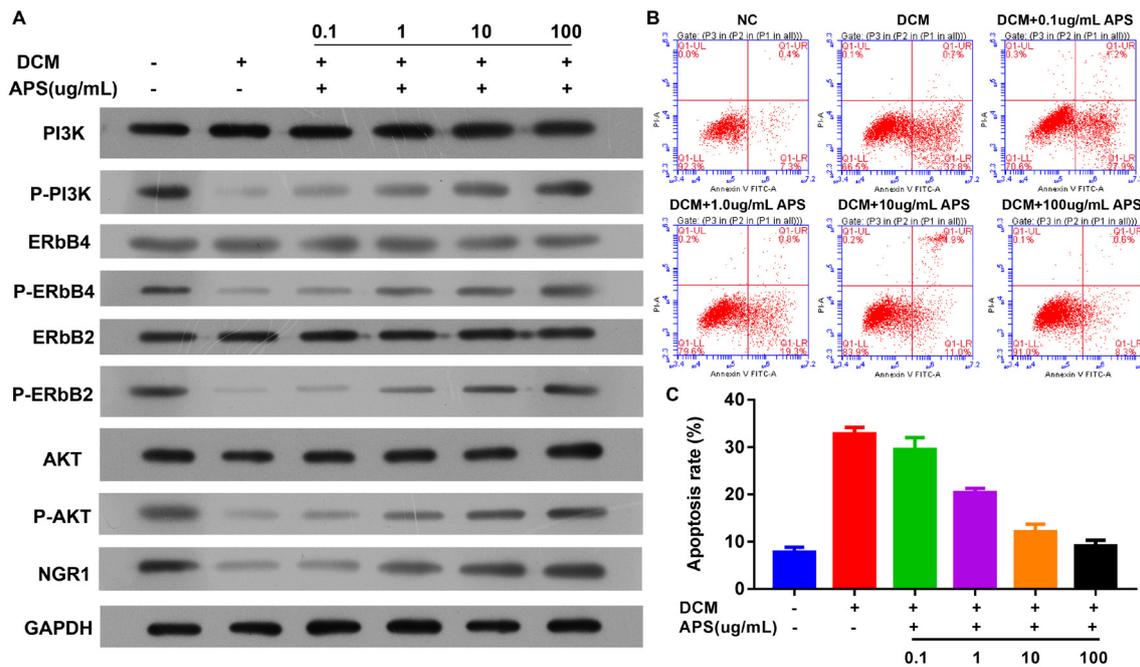


Figure 2. APS inhibits apoptosis and activates NRG1/ErbB pathway. (A) The effect of APS on NRG1/ErbB pathway (level of NRG1, ErbB2 and ErbB4 and phosphorylation level of ErbB2 and ErbB4) and downstream AKT/PI3K signaling (expression level of AKT, PI3K and phosphorylation level of AKT and PI3K) in DCM model H9C2 cells is examined by western blot. GAPDH expression was used as internal control. (B and C) The function of APS on apoptosis in DCM model H9C2 cells is analyzed by flow cytometry. ** $p < 0.01$ vs. DCM group cells.

Activity of GSH-Px and SOD, and level of MDA and NO are markers of oxidative stress. Our data revealed, compared with the NC group cells, the activity of GSH-Px and SOD was decreased while MDA and NO levels were increased in DCM model cells (Figure 1C, 1D, 1F and 1E) which also proved DCM is correlated with oxidative stress. Conversely, APS inhibited activity of GSH-Px and SOD and increased the level of MDA and NO, indicating APS may exert therapeutic effect for DCM by antioxidative function.

3.2. APS inhibits apoptosis and activates NRG1/ErbB pathway

Previous studies showed, NRG1 can bind to ErbB receptor to downstream signaling effectors such as AKT/PI3K pathway, thereby, NRG1/ErbB signaling was involved in various biology processes. It proved that NRG1 increased oxidative capacity and improved mitochondrial function. Besides, NRG1/ErbB pathway was impaired in DCM suggesting this pathway may be involved in pathogenesis of DCM. Here, in this study, western blot results showed APS increased expression level of NGR1 in DCM model cells dose-dependently (Figure 2A). APS did not affect expression level of ErbB2 or ErbB4, whereas the phosphorylation level of ErbB2 and ErbB4 was increased (Figure 2A). In addition, APS-induced NRG1 activated downstream AKT/PI3K pathway which promoted phosphorylation of AKT and PI3K (Figure 2A). NRG1/ErbB and

downstream AKT/PI3K pathway are involved in apoptosis. Besides, suppression on apoptosis cardiac cells is one of the therapeutic strategies for DCM. Our data proved, compared with NC group cells, apoptosis was higher in DCM model cells (Figure 2B and 2C); and, APS exerted inhibitory function on apoptosis in DCM model cells dose-dependently (Figure 2B and 2C). Our results suggested APS may exert inhibiting apoptosis by activating NRG1/ErbB pathway.

3.3. Canertinib partly abolishes the function of APS on proliferation and antioxidation

Our results suggested APS may exert its function by activating NRG1/ErbB pathway. Therefore, ErbB inhibitor (Canertinib) was used as the tool to explore the mechanism of APS in DCM model H9C2 cells. DCM model H9C2 cells were incubated with 100 $\mu\text{g}/\text{mL}$ APS and/or Canertinib. Then, the cell viabilities were tested by CCK-8 assay. As previous showed, APS promoted proliferation whereas Canertinib can partly reverse this promotional function (Figure 3A). Besides, Canertinib can also partly abolish the inhibitory effect of APS on intracellular ROS level (Figure 3B). Previously, We have showed APS suppressed the activity of GSH-Px and SOD and increased the level of MDA and NO (Figure 1C, 1D, 1E and 1F). Here, the results showed Canertinib treatment can partly reverse the effect of APS on these oxidative stress markers (Figure 3C, 3D, 3E and 3F). Taken together, our data proved APS promoted

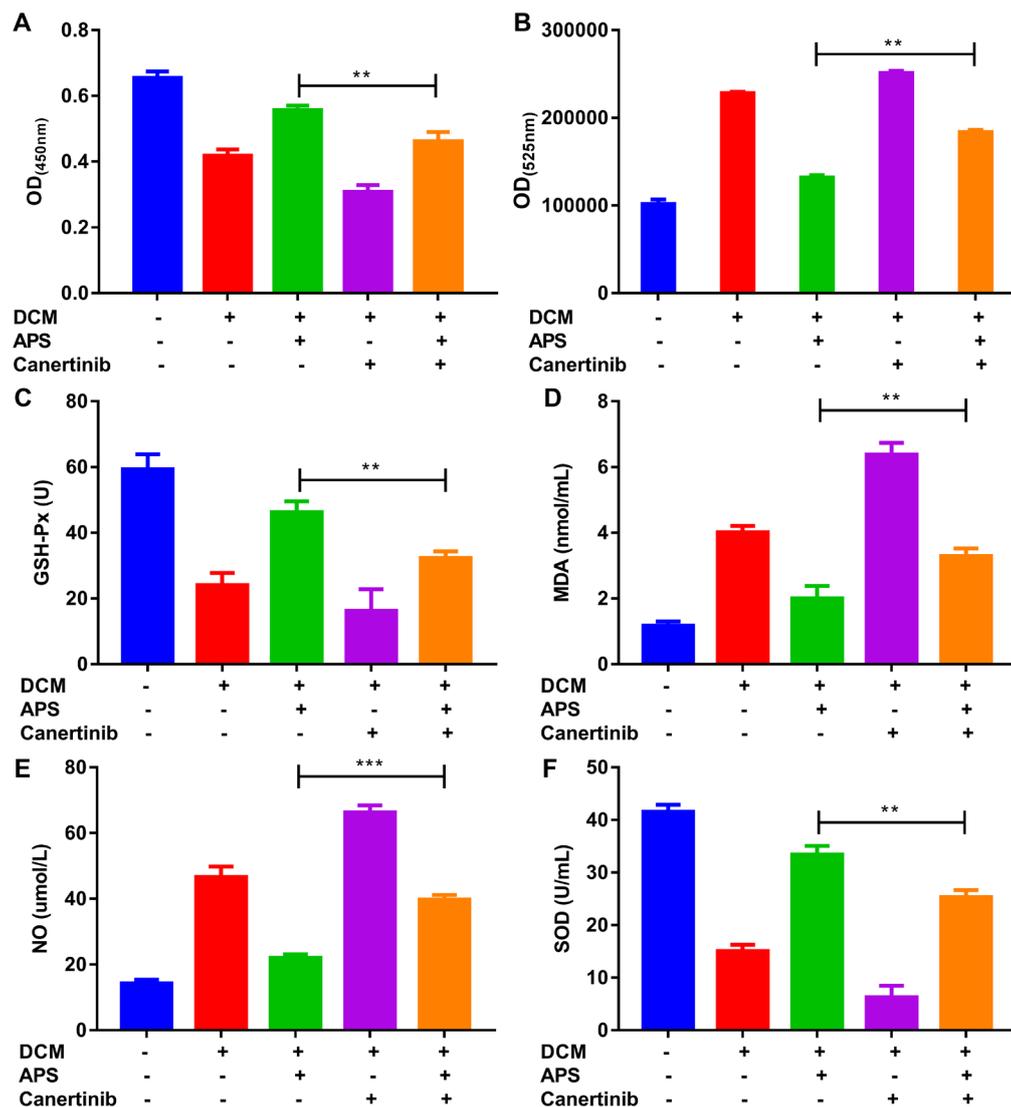


Figure 3. Canertinib partly abolishes the function of APS on proliferation and antioxidation. (A and B) DCM model H9C2 cells were incubated with 100 μ g/mL APS and/or Canertinib for 48 hours. Then, the cell viabilities of the treated was tested by CCK-8 assay (A) and intracellular ROS level was tested by DCF (B). (C, D, E, and F) The DCM model H9C2 cells were treated as A and B, then, the activity of GSH-Px (C) and SOD (F), and the levels of MDA (D) and NO (E) were detected by ELISA assay. Data was presented as Mean \pm SD, * p < 0.05, ** p < 0.01 and *** p < 0.001, DCM + APS vs. DCM + APS + Canertinib.

proliferation and antioxidation in DCM model cells by activating ErbB.

3.4. APS suppresses apoptosis in DCM model H9C2 cells by activating NRG1/ErbB pathway

Previously, we showed APS inhibited apoptosis in DCM model H9C2 cells (Figure 2B). To explore whether ErbB was associated with the effect of APS on apoptosis in DCM model H9C2 cells, DCM model cells were treated with APS and/or canertinib, then, apoptosis was analyzed by flow cytometry. The analysis revealed, canertinib can partly reverse the inhibitory function of APS on apoptosis in DCM model cells (Figure 4A and 4B). This result proved ErbB was involved in the function of APS on apoptosis. APS-induced NRG1 can activate ErbB and the downstream

AKT/PI3K signaling (Figure 2A), while canertinib can partly restore APS-induced NRG1 expression (Figure 4C). Further analyze showed canertinib also reversed APS-induced phosphorylation level of ErbB2/4, AKT and PI3K (Figure 4C). In summary, our data proved APS exerted its protective function for DCM model cells by activating NRG1/ErbB pathway.

4. Discussion

DCM is one of the most common complications of DM. DCM is responsible for almost 80% of the mortality of the diabetic patients. Initially, DCM was defined as the results of coronary artery disease and abnormal myocardial structure (3-6). However, recently, the studies revealed DCM is a result of the long persistent process of the metabolic effect of DM on myocardium.

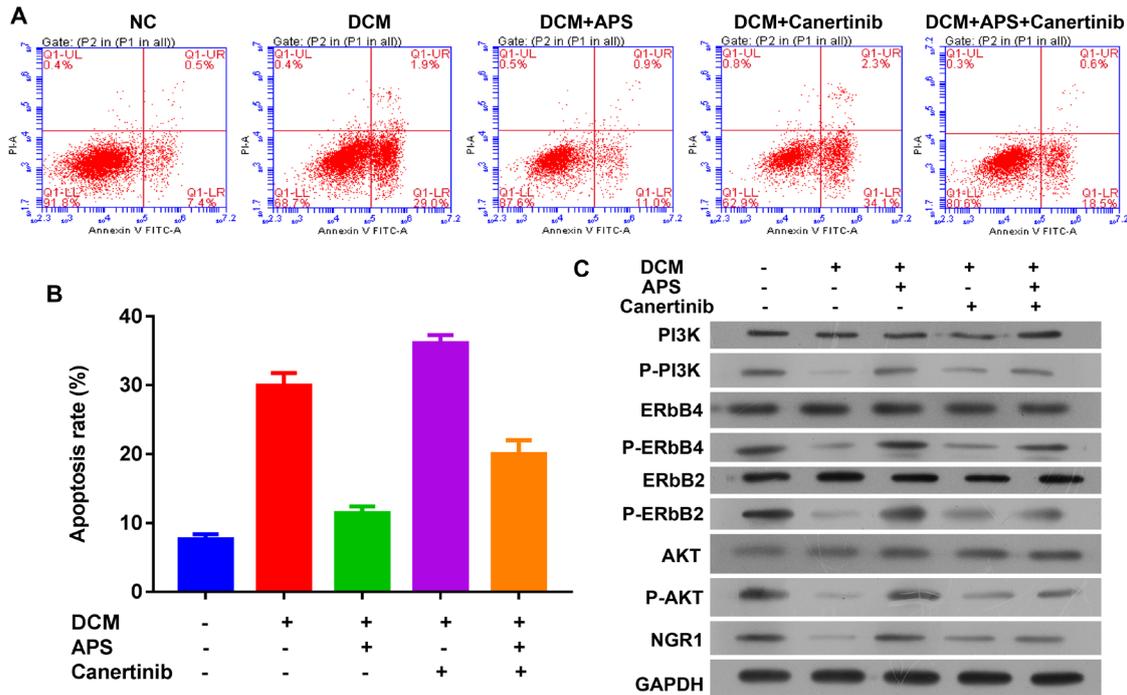


Figure 4. APS suppresses apoptosis in DCM model H9C2 cells by activating NRG1/ErbB pathway. (A and B) DCM model cells were incubated with APS and/or Canertinib for 48 hours, then, apoptosis was analyzed by flow cytometry. **(C)** DCM model cells were treated with APS and/or Canertinib. After 48 hours, the samples were harvested and the effect of canertinib on APS-activated NRG1/ErbB pathway (expression of NRG1, ErbB2, ErbB4 and phosphorylation level of ErbB2 and ErbB4) and downstream AKT/PI3K signaling (expression level of AKT, PI3K and phosphorylation level of AKT and PI3K) was tested by western blot. GAPDH expression was used as internal control.

And, the pathology of DCM is independent of hypertension and coronary. But, the pathogenesis of DCM is not fully understood right now. According to previous studies, multiple biological processes are involved in the pathogenesis and progression of DCM, such as cardiomyocyte apoptosis, endoplasmic reticulum, autophagy, mitochondrial dysfunction (7-10). Among these various biology processes, oxidative stress is the most important process inducing DM to DCM (11-16). Reactive oxygen species (ROS) are the most common chemically reactive chemical species, it showed that over accumulated ROS can cause oxidative stress (11-16). Here, in this study, our data also showed, compared with the normal control (NC) cells, intracellular ROS level is higher, further studies displayed the activity of GSH-Px was decreased while the levels of MDA and NO were elevated in DCM model cells (Figure 1C, 1D, 1E and 1F). Our results proved the previous studies in turn. Therefore, inhibiting oxidative stress is believed to be an important strategy for the therapy of DCM. Previous studies revealed, NRG1/ErbB pathway is impaired in DCM cells which suggested NRG1/ErbB pathway may play an important role in DCM (30-32). Besides, NRG1/ErbB can improve glucose tolerance in healthy and diabetic rodents (32). And, also previous studies showed NRG1 can regulate myocyte oxidative capacity (33,34). Our data revealed, compared with the normal control cells, NRG1/ErbB pathway is inactive in DCM model cells (Figure 2A).

Recently, increasing evidences showed natural products exerted antioxidative effect in DCM. APS is one of the key active component of the traditional Chinese medical herb *Astragalus membranaceus* (20,22-24,26,35). Previously, it's showed that APS decreased apoptosis in high-glucose-induced H9C2 cells by regulating the function of mitochondria and inhibiting expressing of caspase (23). Present, we showed that, APS improved proliferation (Figure 1A) and decreased apoptosis (Figure 2B and 2C) in DCM model H9C2 cells. Further studies showed, firstly, APS lowered the intracellular ROS level in DCM model cells (Figure 1B). Secondly, APS can also increase the activity of GSH-Px and SOD and decreased the level of MDA and NO in DCM model H9C2 cells (Figure 1C, 1D, 1E and 1F). These results indicated APS may play protective role in DCM cells. Our results proved, compared with the normal control cells, NRG1/ErbB pathway was inactive in DCM model H9C2 cells (Figure 2A). And, APS activated the NRG1/ErbB pathway dose-dependently and time-dependently (Figure 2A), which suggested APS may exert its protective role in DCM model H9C2 cells by activating NRG1/ErbB pathway. Canertinib is the inhibitor of ErbB. Pretreatment of DCM model cells with APS and Canertinib can partly abolish the protective effect of APS in DCM model H9C2 cells by inhibiting NRG1/ErbB pathway. Canertinib can partly reverse APS-induced proliferation (Figure 3A). And, canertinib can also partly abrogate the function of APS

on apoptosis (Figure 4A), intracellular ROS level (Figure 3B) and the activity of GSH-Px and SOD and the levels of MDA and NO (Figure 3C, 3D, 3E and 3F). Taken together, our results proved APS exerted its protective role by activating NRG1/ErbB pathway.

In conclusion, our study proved: *i*) DCM is associated with proliferation, apoptosis, oxidative stress and inactive NRG1/ErbB pathway; *ii*) APS can increase proliferation, inhibit apoptosis and improve antioxidative function including reducing intracellular ROS level, elevating activity of GSH-Px and SOD and lowering the level of MDA and NO by activating NRG1/ErbB pathway. Our study broadened the mechanisms of DCM and increased potential for prevention and therapy to DCM.

Acknowledgements

This work was supported by National Natural Science foundation of China (NO: 81774182), Science and Technology Program of Guangzhou (NO. 201607010367) and Medical Scientific Research Foundation of Guangdong Province (NO. A2016231).

References

- Ogurtsova K, da Rocha Fernandes JD, Huang Y, Linnenkamp U, Guariguata L, Cho NH, Cavan D, Shaw JE, Makaroff LE. IDF Diabetes Atlas: Global estimates for the prevalence of diabetes for 2015 and 2040. *Diabetes Res Clin Pract.* 2017; 128:40-50.
- Guariguata L, Whiting DR, Hambleton I, Beagley J, Linnenkamp U, Shaw JE. Global estimates of diabetes prevalence for 2013 and projections for 2035. *Diabetes Res Clin Pract.* 2014; 103:137-149.
- Waddingham MT, Edgley AJ, Tsuchimochi H, Kelly DJ, Shirai M, Pearson JT. Contractile apparatus dysfunction early in the pathophysiology of diabetic cardiomyopathy. *World J Diabetes.* 2015; 6:943-960.
- Trachanas K, Sideris S, Aggeli C, Poulidakis E, Gatzoulis K, Tousoulis D, Kallikazaros I. Diabetic cardiomyopathy: From pathophysiology to treatment. *Hellenic J Cardiol.* 2014; 55:411-421.
- Pappachan JM, Varughese GI, Sriraman R, Arunagirinathan G. Diabetic cardiomyopathy: Pathophysiology, diagnostic evaluation and management. *World J Diabetes.* 2013; 4:177-189.
- Miki T, Yuda S, Kouzu H, Miura T. Diabetic cardiomyopathy: Pathophysiology and clinical features. *Heart Fail Rev.* 2013; 18:149-166.
- Guo R, Wu Z, Jiang J, Liu C, Wu B, Li X, Li T, Mo H, He S, Li S, Yan H, Huang R, You Q, Wu K. New mechanism of lipotoxicity in diabetic cardiomyopathy: Deficiency of Endogenous H₂S Production and ER stress. *Mech Ageing Dev.* 2017; 162:46-52.
- Wang K, Han D, Zhang Y, Rong C, Zhang Y. Protective effect and mechanism of beta-CM7 on renin angiotensin system & diabetic cardiomyopathy. *Sheng Wu Gong Cheng Xue Bao.* 2016; 32:195-203. (in Chinese)
- Ni R, Zheng D, Xiong S, Hill DJ, Sun T, Gardiner RB, Fan GC, Lu Y, Abel ED, Greer PA, Peng T. Mitochondrial calpain-1 disrupts ATP synthase and induces superoxide generation in type 1 diabetic hearts: A novel mechanism contributing to diabetic cardiomyopathy. *Diabetes.* 2016; 65:255-268.
- Liu JW, Liu D, Cui KZ, Xu Y, Li YB, Sun YM, Su Y. Recent advances in understanding the biochemical and molecular mechanism of diabetic cardiomyopathy. *Biochem Biophys Res Commun.* 2012; 427:441-443.
- Aksakal E, Akaras N, Kurt M, Tanboga IH, Halici Z, Odabasoglu F, Bakirci EM, Unal B. The role of oxidative stress in diabetic cardiomyopathy: An experimental study. *Eur Rev Med Pharmacol Sci.* 2011; 15:1241-1246.
- Watanabe K, Thandavarayan RA, Harima M, Sari FR, Gurusamy N, Veeraveedu PT, Mito S, Arozal W, Sukumaran V, Laksmanan AP, Soetikno V, Kodama M, Aizawa Y. Role of differential signaling pathways and oxidative stress in diabetic cardiomyopathy. *Curr Cardiol Rev.* 2010; 6:280-290.
- Umbarkar P, Singh S, Arkat S, Bodhankar SL, Lohidasan S, Sitasawad SL. Monoamine oxidase-A is an important source of oxidative stress and promotes cardiac dysfunction, apoptosis, and fibrosis in diabetic cardiomyopathy. *Free Radic Biol Med.* 2015; 87:263-273.
- Yang R, Jia Q, Liu XF, Gao Q, Wang L, Ma SF. Effect of hydrogen sulfide on oxidative stress and endoplasmic reticulum stress in diabetic cardiomyopathy. *Zhongguo Ying Yong Sheng Li Xue Za Zhi.* 2016; 32:8-12. (in Chinese)
- Thandavarayan RA, Giridharan VV, Watanabe K, Konishi T. Diabetic cardiomyopathy and oxidative stress: Role of antioxidants. *Cardiovasc Hematol Agents Med Chem.* 2011; 9:225-230.
- Liu Q, Wang S, Cai L. Diabetic cardiomyopathy and its mechanisms: Role of oxidative stress and damage. *J Diabetes Investig.* 2014; 5:623-634.
- Alonso N, Moliner P, Mauricio D. Pathogenesis, clinical features and treatment of diabetic cardiomyopathy. *Adv Exp Med Biol.* 2017. doi: 10.1007/5584_2017_105.
- Huynh K, Bernardo BC, McMullen JR, Ritchie RH. Diabetic cardiomyopathy: Mechanisms and new treatment strategies targeting antioxidant signaling pathways. *Pharmacol Ther.* 2014; 142:375-415.
- Falcao-Pires I, Leite-Moreira AF. Diabetic cardiomyopathy: Understanding the molecular and cellular basis to progress in diagnosis and treatment. *Heart Fail Rev.* 2012; 17:325-344.
- Chen W, Yu MH, Li YM, Chen WJ, Xia YP. Beneficial effects of astragalus polysaccharides treatment on cardiac chymase activities and cardiomyopathy in diabetic hamsters. *Acta Diabetol.* 2010; 47 Suppl 1:35-46.
- Hayat SA, Patel B, Khattar RS, Malik RA. Diabetic cardiomyopathy: Mechanisms, diagnosis and treatment. *Clin Sci (Lond).* 2004; 107:539-557.
- Ju J, Chen W, Lai Y, Wang L, Wang H, Chen WJ, Zhao X, Ye H, Li Y, Zhang Y. Astragalus polysaccharides improve cardiomyopathy in STZ-induced diabetic mice and heterozygous (SOD2^{+/-}) knockout mice. *Braz J Med Biol Res.* 2017; 50:e6204.
- Sun S, Yang S, Dai M, Jia X, Wang Q, Zhang Z, Mao Y. The effect of Astragalus polysaccharides on attenuation of diabetic cardiomyopathy through inhibiting the extrinsic and intrinsic apoptotic pathways in high glucose-stimulated H9C2 cells. *BMC Complement Altern Med.* 2017; 17:310.
- Chen W, Li YM, Yu MH. Astragalus polysaccharides inhibited diabetic cardiomyopathy in hamsters depending

- on suppression of heart chymase activation. *J Diabetes Complications*. 2010; 24:199-208.
25. Cui K, Zhang S, Jiang X, Xie W. Novel synergic antidiabetic effects of Astragalus polysaccharides combined with Crataegus flavonoids *via* improvement of islet function and liver metabolism. *Mol Med Rep*. 2016; 13:4737-4744.
 26. Chen W, Xia Y, Zhao X, Wang H, Chen W, Yu M, Li Y, Ye H, Zhang Y. The critical role of Astragalus polysaccharides for the improvement of PPARalpha [correction of PPRAalpha]-mediated lipotoxicity in diabetic cardiomyopathy. *PLoS One*. 2012; 7:e45541.
 27. Chen W, Xia YP, Chen WJ, Yu MH, Li YM, Ye HY. Improvement of myocardial glycolipid metabolic disorder in diabetic hamster with Astragalus polysaccharides treatment. *Mol Biol Rep*. 2012; 39:7609-7615.
 28. Tian J, Zhao Y, Liu Y, Liu Y, Chen K, Lyu S. Roles and mechanisms of herbal medicine for diabetic cardiomyopathy: Current status and perspective. *Oxid Med Cell Longev*. 2017; 2017:8214541.
 29. Wassef MAE, Tork OM, Rashed LA, Ibrahim W, Morsi H, Rabie DMM. Mitochondrial dysfunction in diabetic cardiomyopathy: Effect of mesenchymal stem cell with PPAR-gamma agonist or exendin-4. *Exp Clin Endocrinol Diabetes*. 2018; 126:27-38.
 30. Gui C, Zhu L, Hu M, Lei L, Long Q. Neuregulin-1/ ErbB signaling is impaired in the rat model of diabetic cardiomyopathy. *Cardiovasc Pathol*. 2012; 21:414-420.
 31. Li B, Zheng Z, Wei Y, Wang M, Peng J, Kang T, Huang X, Xiao J, Li Y, Li Z. Therapeutic effects of neuregulin-1 in diabetic cardiomyopathy rats. *Cardiovasc Diabetol*. 2011; 10:69.
 32. Ennequin G, Capel F, Caillaud K, Chavanelle V, Etienne M, Teixeira A, Li X, Boisseau N, Sirvent P. Neuregulin 1 improves complex 2-mediated mitochondrial respiration in skeletal muscle of healthy and diabetic mice. *Sci Rep*. 2017; 7:1742.
 33. Odiete O, Hill MF, Sawyer DB. Neuregulin in cardiovascular development and disease. *Circ Res*. 2012; 111:1376-1385.
 34. Rupert CE, Coulombe KL. The roles of neuregulin-1 in cardiac development, homeostasis, and disease. *Biomark Insights*. 2015; 10(Suppl 1):1-9.
 35. Chen W, Li YM, Yu MH. Effects of Astragalus polysaccharides on chymase, angiotensin-converting enzyme and angiotensin II in diabetic cardiomyopathy in hamsters. *J Int Med Res*. 2007; 35:873-877.

(Received February 10, 2018; Revised March 9, 2018; Accepted March 20, 2018)

Distinct pattern of Th17/Treg cells in pregnant women with a history of unexplained recurrent spontaneous abortion

Jinfeng Qian^{1,2,3,§}, Na Zhang^{1,2,3,§}, Jing Lin^{1,2,3,§}, Caiyan Wang^{1,2,3}, Xinyao Pan^{1,2,3}, Lanting Chen^{1,2,3}, Dajin Li^{1,2,3,*}, Ling Wang^{1,2,3,*}

¹Laboratory for Reproductive Immunology, Hospital & Institute of Obstetrics and Gynecology, Shanghai Medical College, Fudan University, Shanghai, China;

²The Academy of Integrative Medicine of Fudan University, Shanghai, China;

³Shanghai Key Laboratory of Female Reproductive Endocrine Related Diseases, Shanghai, China.

Summary

The aim of the current study was to determine the pattern of immune cells and related functional molecules in peripheral blood and at the maternal-fetal interface in women with unexplained recurrent spontaneous abortion (URSA). In part I, 155 women were included and divided into four groups: non-pregnant controls with no history of URSA (NPCs), pregnant controls with no history of URSA (PCs), non-pregnant women with a history of URSA (NPU), and pregnant women with a history of URSA (PU). Venous blood samples were collected and analyzed. In part II, 35 subjects with URSA and 40 subjects in the early stage of normal pregnancy who chose to undergo an abortion were recruited. Samples of the decidua were collected, and the proportion of immune cells and the expression of related molecules were evaluated. Peripheral regulatory T cells (Treg cells) increased in PCs compared to NPCs, but in women with URSA the flux of Treg cells disappeared when pregnancy occurred. Levels of interleukin-10 (IL-10), cytotoxic T lymphocyte-associated antigen 4 (CTLA-4), and IL-17 and the ratio of Th17/Treg cells in peripheral blood remained stable among the four groups. At the maternal-fetal interface, the percentage of Treg cells, the level of CTLA-4 of CD4⁺CD25⁺CD127^{lo} cells and CD4⁺Foxp3⁺ cells were significantly lower in women with URSA compared to controls, respectively. Levels of transforming growth factor-β1 (TGF-β1) mRNA and protein in the decidua significantly decreased in URSA while levels of IL-6 and tumor necrosis factor-α (TNF-α) and the Th17/Treg ratio significantly increased. In conclusion, peripheral Treg cells did not increase in pregnant women with URSA. The decrease in Treg cells and levels of CTLA-4 and TGF-β1 and as well as the increase in levels of IL-6 and TNF-α, and the Th17/Treg ratio at the maternal-fetal interface might contribute to inappropriate maternal-fetal immune tolerance in URSA.

Keywords: Treg cells, Th17 cells, unexplained recurrent spontaneous abortion, cytokine, maternal-fetal tolerance

Released online in J-STAGE as advance publication April 15, 2018.

§These authors contributed equally to this work.

*Address correspondence to:

Dr. Dajin Li, Laboratory for Reproductive Immunology, Hospital & Institute of Obstetrics and Gynecology, Fudan University, 419 Fangxie Road, Shanghai, 200011, China.
E-mail: djli@shmu.edu.cn

Dr. Ling Wang, Laboratory for Reproductive Immunology, Hospital & Institute of Obstetrics and Gynecology, Fudan University, 419 Fangxie Road, Shanghai, 200011, China.
E-mail: Dr.wangling@fudan.edu.cn

1. Introduction

Recurrent spontaneous abortion (RSA) refers to two or more consecutive pregnancy losses before 20 weeks of gestation. RSA affects about 1% of all women and results in physical and psychological distress. The common causes of RSA include parental or embryonic karyotype anomalies, uterine anatomic abnormalities, infection, endocrine disorders, and antiphospholipid syndrome (1). However, nearly 50% of cases remain unexplained (2). Maternal-fetal immune abnormalities

might be one cause of unexplained recurrent spontaneous abortion (URSA) (3).

The human fetus is a semi-allograft and antigenically foreign to the mother. Maternal-fetal immune tolerance plays a significant role in establishing and maintaining a successful pregnancy. CD4⁺ T cells, also known as Th cells, are crucial to this process. CD4⁺ T cells can be classified into Th1 cells, Th2 cells, regulatory T cells (Treg cells), and Th17 cells (4). Previously, a successful pregnancy was believed to be associated with Th2-dominant immunity, and an imbalance in Th1/Th2 cells was believed to result in spontaneous abortion (5). However, Th2-dominance was later found in URSA (6). Therefore, the Th1/Th2 paradigm was no longer sufficient to explain the maternal-fetal tolerance, and a new paradigm of Th1/Th2/Th17/Treg cells has been developed.

Treg and Th17 cells share the same origin and have the opposite effects on inflammation. Treg cells express surface markers such as CD25 (7), CD127 (low expression, human only) (8), cytotoxic T lymphocyte-associated antigen 4 (CTLA-4) (9), and glucocorticoid induced tumor necrosis factor receptor-related protein (GITR) (10). Forkhead box protein 3 (Foxp3) serves as a unique intracellular transcription factor for the development of Treg cells (11). The main cytokines secreted by Treg cells include interleukin-10 (IL-10) and transforming growth factor β (TGF- β). Treg cells exhibit anti-inflammatory and immune-suppressive action mainly through cytokines and contact-dependent suppression (12). Th17 cells play an important role in autoimmunity and the protective response against extracellular fungi and bacteria by releasing pro-inflammatory cytokines such as IL17, IL-6, IL-22, and tumor necrosis factor- α (TNF- α) (13). Th17 cells are regulated by transcription factor retinoic acid receptor-related orphan receptors (ROR) (14).

The essential role of Treg cells in growth and development of the fetus in humans and mice has been widely reported (15). The function of Th17 cells in maternal-fetal tolerance has not been fully elucidated. Changes in the proportion of Th17 cells in women in the early stage of normal pregnancy compared to the proportion in non-pregnant women are not consistent (16). Recent findings indicated that Th17 cells might participate in pregnancy-related complications including spontaneous abortions and pre-eclampsia (17). The Th17/Treg ratio shifts in favor of Treg cells in a health pregnancy, and a decrease in Treg cells or an increase in Th17 cells is detrimental to normal pregnancy (18). However, the roles of Th17 cells and the balance in Th17/Treg cells in the early stage of pregnancy with URSA are still unclear.

The current study focused on the pattern of Th17/Treg cells and the expression of related molecules in pregnant women with a history of URSA in peripheral blood as well as at the maternal-fetal interface.

2. Materials and Methods

2.1. Study population

Study subjects ages 18 to 40 years were recruited from the Obstetrics and Gynecology Hospital of Fudan University in China between August 2017 and March 2018. To investigate the peripheral Th17/Treg balance in part I, a total of 155 women were divided into four groups and venous blood samples were collected: 40 non-pregnant controls with no history of URSA (NPCs), 40 pregnant controls with no history of URSA (PCs), 38 non-pregnant women with a history of URSA (NPU), 37 pregnant women with a history of URSA (PU). To study the immune tolerance at the maternal-fetal interface in part II, 35 subjects with URSA and 40 subjects in the early stage of normal pregnancy who chose to undergo an abortion were recruited, and samples of the aborted tissue were collected. Women with a history of two or more consecutive spontaneous abortions due to unexplained causes (excluding parental or embryonic karyotype anomalies, uterine anatomic abnormalities, an infection, endocrine disorders, and antiphospholipid syndrome) before 20 weeks of gestation were diagnosed as having URSA. Women with at least one successful pregnancy or no history of spontaneous abortions served as controls. In pregnant women, the gestational age was from 45 to 90 days when samples were collected. None of the subjects had a history of smoking, alcohol addiction, liver or renal diseases, or vaccination in three months before recruitment. This study was approved by the ethics committee of the Obstetrics and Gynecology Hospital of Fudan University. All subjects provided written informed consent.

2.2. Blood and tissue samples

Peripheral blood mononuclear cells (PBMCs) were isolated from 5 mL of heparinized venous blood by Ficoll-Hypaque density gradient centrifugation (2,500 rpm at room temperature for 20 minutes). Cells were collected at the interface and washed with phosphate buffered saline (PBS). The decidua was aseptically collected within 5 minutes after artificial abortion and washed with sterile saline. One sample was stored at -80°C for a real-time quantitative polymerase chain reaction (RT-PCR) and an enzyme-linked immunosorbent assay (ELISA), and another sample was placed in DMEM/F12 culture for cell separation. In brief, tissue was digested with collagenase and DNase at 37°C for 30 minutes in a shaking water bath. The released cells were then separated by filtering through a nylon mesh, and the cells were centrifuged at 1,500 rpm for 8 minutes at room temperature. Percoll was added to the mixture, which was centrifuged at 2,500 rpm for 30 minutes. The cells were stratified into a dead cell layer,

a decidual stromal cell layer, a decidual immune cell layer, and an erythrocyte layer. The decidual immune cell layer was preserved, washed with PBS, and then centrifuged at 1,500 rpm (4°C) for 8 minutes to obtain a precipitate.

2.3. Flow cytometry analysis

Collected cells were divided into stimulated and unstimulated samples. Surface chemokine data were obtained from the unstimulated samples. CD4, CD25, CD127, and CTLA-4/PEcy7 were used to perform surface staining for 30 minutes in the dark. To stain the intracellular molecule Foxp3, cells were permeabilized with a permeabilization/fixation buffer and stained with antihuman Foxp3-PE antibody followed by surface staining. The stimulated samples were used to stain the intracellular molecules IL-17A and IL-10. Cells were permeabilized with a permeabilization/fixation buffer and stained with antihuman IL17A-Alexa 647 and anti-human IL-10-Alexa 647 antibody (Biolegend), followed by surface staining. CyAn ADP flow cytometer (Beckman Coulter) was used to analyze the samples. Postacquisition FACS data were analyzed using the program Flowjo (Ashland, OR, USA).

2.4. RT-PCR analysis

Total RNA was isolated from decidual tissue. Briefly, 1 mL of Trizol per 100 mg of tissue was added to tissue and the two were mixed thoroughly on ice. After incubation at room temperature for 10 minutes, 200 µL of chloroform was added and the mixture was centrifuged at 12,000 rpm for 15 minutes. The upper aqueous phase containing RNA was transferred to another 1.5-mL sterile RNase-free EP tube and precipitated by adding an equal volume of 100% isopropanol. The RNA was concentrated by centrifugation at 12,000 rpm for 10 minutes (4°C) and washed with 75% ethanol twice. The resulting RNA pellet was then air-dried and dissolved in 30 µL RNase-free water. A reverse transcription reaction was performed at a final volume of 20 µL containing 500 ng of the RNA sample and 2 µL of 5×PrimeScript RT Master Mix (Perfect Real Time). Afterwards, quantitative real-time PCR was performed on a 20-µL mixture of 2 µL of reverse transcription product, 1.6 µL of a forward/reverse primer mixture for a final concentration of 0.4 µM, 10 µL of 2×SYBR Premix Ex TaqII (Tli RNaseH Plus), and DNase-free water added to reach the final volume. Reactions were performed using the following protocol: 95°C for 30 seconds and cycles of 95°C for 5 seconds and 60°C for 30 seconds for 40 cycles, followed by a dissociation stage. The primers used are listed in Table S1 (<http://www.biosciencetrends.com/action/getSupplementalData.php?ID=23>). Gene expression was normalized to the level of the house-keeping gene β-actin.

2.5. ELISA testing

Concentrations of IL-6, IL-10, IL-17, IL-23, TGF-β1, and TNF-α were measured with ELISA according to the manufacturer's instructions (human IL-6 ELISA kit, human IL-10 ELISA kit, human IL-17 ELISA kit, human IL-23 ELISA kit, human TGF-β1 ELISA kit, and human TNF-α ELISA kit; Xitang Biotech Co., Ltd., Shanghai, China). Optical density values were read at 450 nm using a microplate reader (Thermo Fisher Scientific, USA).

2.6. Statistical analysis

The Statistical Package for Social Sciences (SPSS) 17.0 was used for statistical analysis. Data are presented as the mean ± SEM. Before analysis, the data were checked for normality. In part I, data differences were analyzed using one-way ANOVA and a Kruskal-Wallis test. In part II, data with a normal distribution were statistically analyzed using the Student's *t*-test, and non-normal data were statistically assessed with the Wilcoxon signed-rank test. *P* < 0.05 was considered statistically significant.

3. Results

3.1. Clinical characteristics

To investigate the peripheral Th17/Treg balance in part I, a total of 155 women were recruited. Seventy-five subjects had a history of URSA with at least two consecutive first trimester abortions (range: 2-5). There were no significant differences among the four groups in terms of age, body mass index (BMI), and gestational age (Table S2, <http://www.biosciencetrends.com/action/getSupplementalData.php?ID=23>). To study the immune tolerance at the maternal-fetal interface in part II, 75 subjects were recruited. There were no significant differences between the two groups in terms of age, BMI, gestational age, and gravidity (Table S3, <http://www.biosciencetrends.com/action/getSupplementalData.php?ID=23>).

3.2. Peripheral Treg cells in pregnant women with URSA

Results indicated that there was no significant difference in the percentage of CD4⁺ T cells in PBMCs among the four groups (*F* = 2.160, *P* = 0.095) (Figure 1 A, B). The percentage of CD4⁺CD25⁺CD127^{lo} Treg cells among CD4⁺ T cells was significantly higher in PCs than that in NPCs (*t* = 3.165, *P* = 0.002) (Figure 1 C, D). However, in women with URSA, there was no elevation of CD4⁺CD25⁺CD127^{lo} Treg cells among CD4⁺ T cells in PUs compared to NPU (*t* = 1.882, *P* = 0.064) (Figure 1 C, D). The proportion of

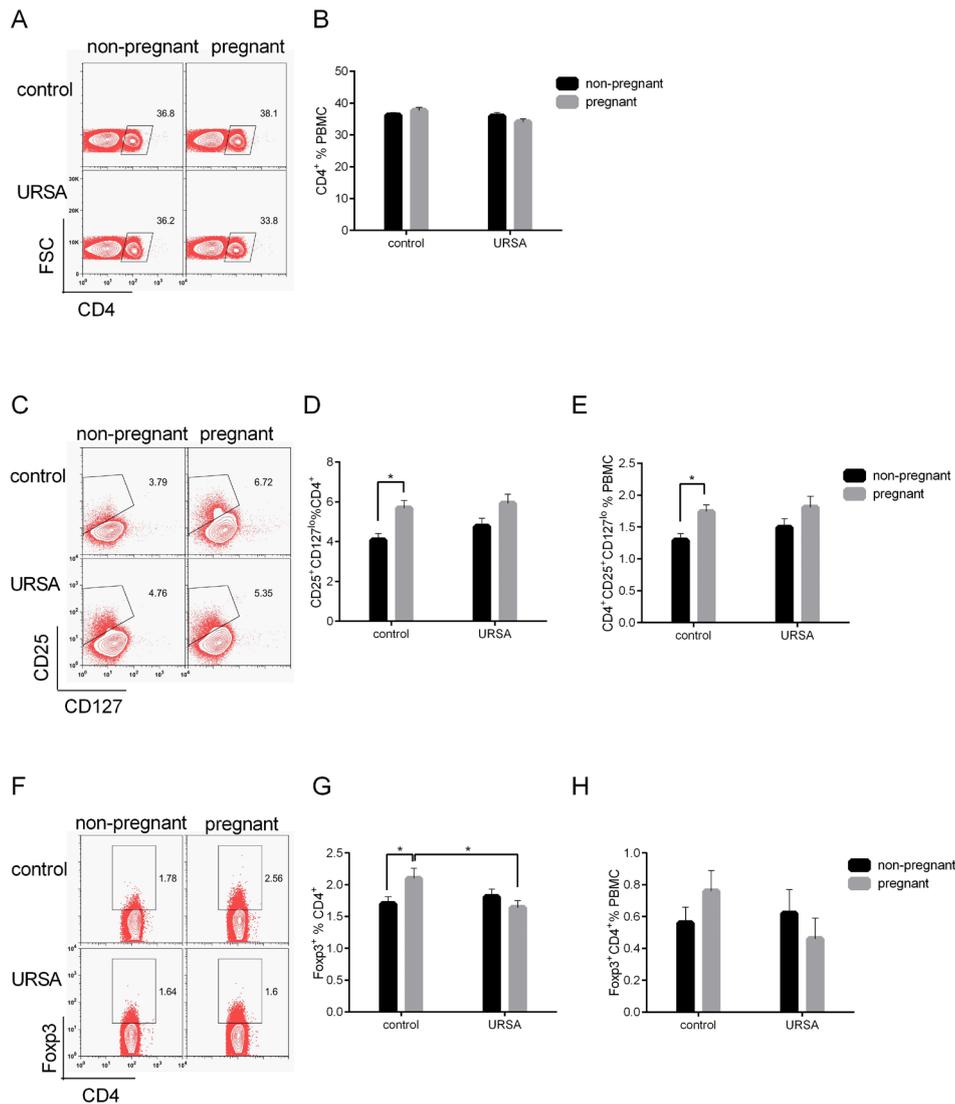


Figure 1. Pattern of peripheral Treg cells in URSA. (A, B) There was no significant difference in the percentage of CD4⁺ T cells in PBMCs among the four groups. (C, D, E) The percentage of CD25⁺CD127^{lo} Treg cells in CD4⁺ T cells increased in pregnant controls (PCs) compared to that in non-pregnant controls (NPCs). CD25⁺CD127^{lo} Treg cells did not increase in pregnant women with a history of URSA (PUs) compared to non-pregnant women with a history of URSA (NPU). Similar results were obtained for CD4⁺CD25⁺CD127^{lo} Treg cells in PBMCs. (F, G, H) The proportion of Foxp3⁺ Treg cells in CD4⁺ T cells was higher in PCs compared to NPCs and lower in PUs than in PCs while remaining stable in PUs compared to NPUs. The proportion of Foxp3⁺ Treg cells in PBMCs did not differ significantly among the four groups. (The FACS picture represents an individual sample. Data represent the mean ± SEM. **p* < 0.05).

CD4⁺CD25⁺CD127^{lo} Treg cells in PBMCs was higher in PCs than that in NPCs (*t* = 2.905, *P* = 0.005) whereas there was no significant difference between that in PUs and that in NPUs (*t* = 1.465, *P* = 0.147) (Figure 1 C, D, E).

Results for Foxp3⁺ Treg cells indicated that the percentage of Treg cells among CD4⁺ T cells increased in PCs compared to that in NPCs (*t* = 2.070, *P* = 0.042) while remaining stable between that in PUs and NPUs (*t* = 1.015, *P* = 0.313) (Figure 1 F, G). The proportion of Foxp3⁺ Treg cells in CD4⁺ T cells was significantly lower in PUs than that in PCs (*t* = 2.430, *P* = 0.018) (Figure 1 F, G). The proportion of Foxp3⁺ Treg cells in PBMCs did not differ significantly among the four groups (Figure 1 F, G, H).

3.3. The expression of IL-10 and CTLA-4 in peripheral blood of women with a history of URSA

The expression of IL-10 and CTLA-4 was studied as a possible pathway for the action of Treg cells. Results indicated that there was no significant difference in the expression of IL-10 by CD4⁺ T cells (*F* = 0.975, *P* = 0.406) (Figure 2 A, B) or IL-10⁺CD4⁺ T cells in PBMCs among the four groups (*F* = 1.043, *P* = 0.376) (Figure 2 C). In addition, there was no significant difference in the expression of CTLA-4 on CD4⁺CD25⁺CD127^{lo} cells among the groups (*F* = 1.301, *P* = 0.276) (Figure 2 D, E).

3.4. The ratio of peripheral Th17/Treg cells in women with a history of URSA

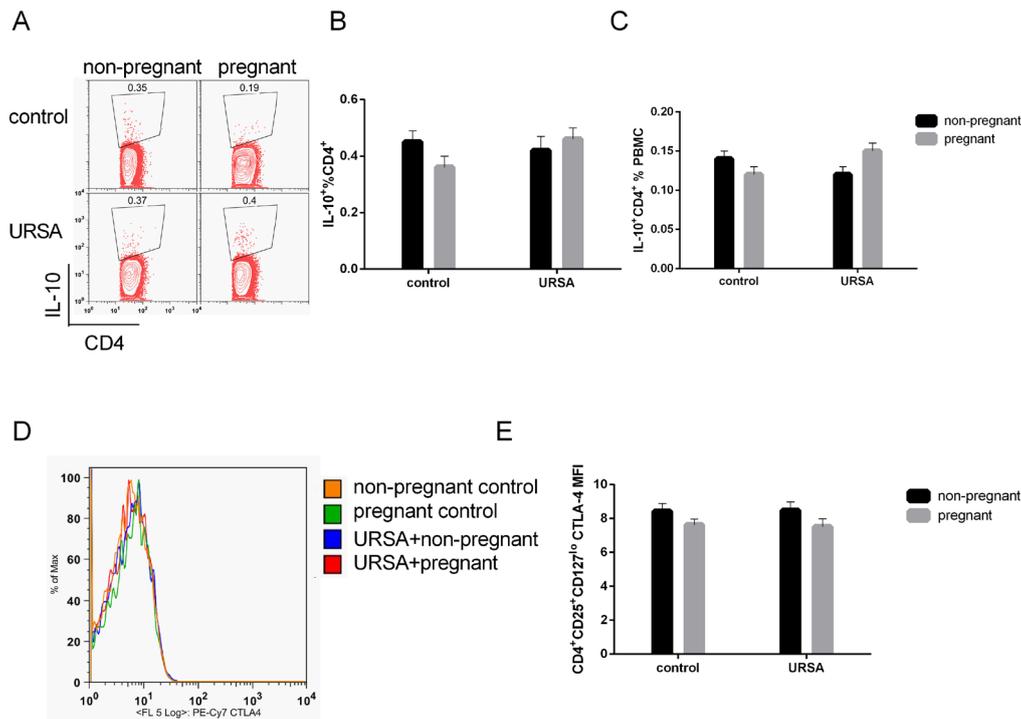


Figure 2. Expression of IL-10 and CTLA-4 in peripheral blood of women with URSA. (A, B, C) There was no significant difference in the expression of IL-10 by CD4⁺ T cells or PBMCs among the four groups. (D, E) There was no significant difference in the CTLA-4 mean fluorescence intensity (MFI) of CD4⁺CD25⁺CD127^{lo} cells among the groups. (The FACS picture represents an individual sample. Data represent the mean \pm SEM. * $p < 0.05$.)

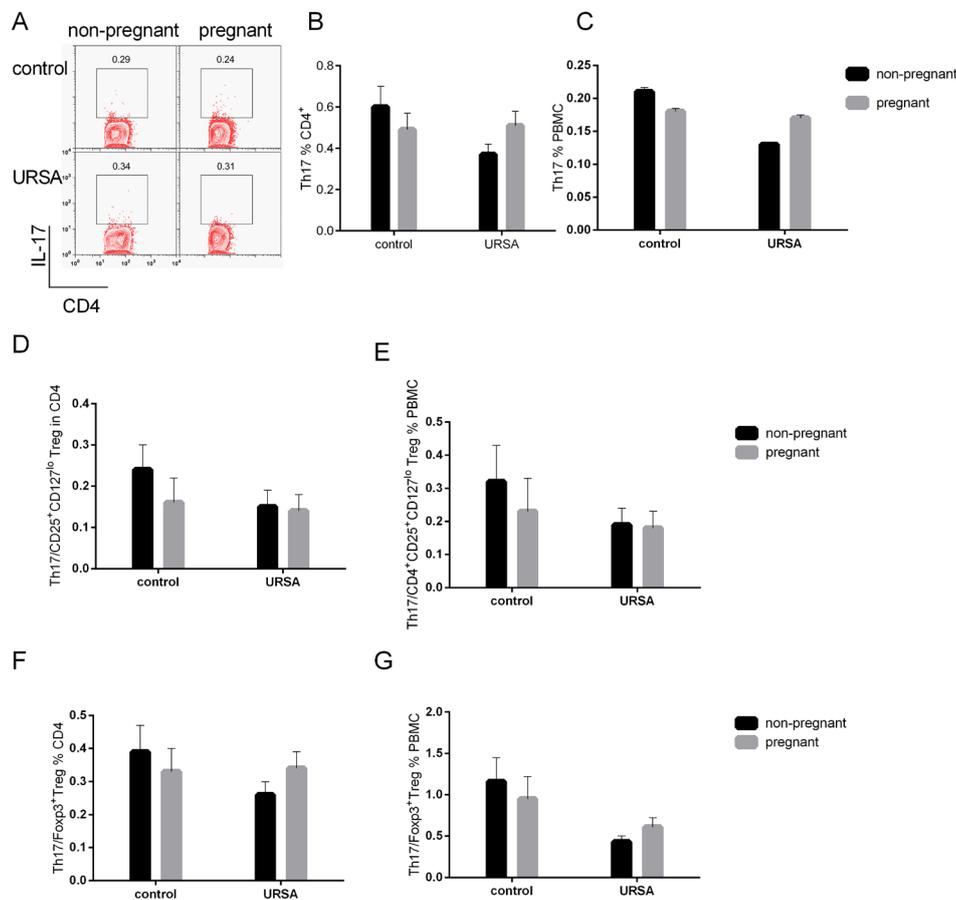


Figure 3. The ratio of peripheral Th17/Treg in URSA. (A, B, C) There was no significant difference in the expression of IL-17 by CD4⁺ T cells or PBMCs among the four groups. (D, E, F, G) There was no significant difference in the ratio of Th17/Treg cells among CD4⁺ T cells or PBMCs among the groups. (The FACS picture represents an individual sample. Data represent the Mean \pm SEM. * $p < 0.05$.)

IL-17 is a characteristic cytokine of Th17 cells. The current study found that IL-17⁺ cells among CD4⁺ T cells or PBMCs remained stable among the four groups ($F = 1.386, P = 0.250$; $F = 1.099, P = 0.352$) (Figure 3 A, B, C). The ratio of Th17/Treg cells was examined, and results indicated that there was no statistical difference in either CD4⁺ T cells (Th17/CD4⁺CD25⁺CD127^{lo} Treg; $Z = 3.305, P = 0.347$; Th17/Foxp3⁺ Treg; $Z = 4.919, P = 0.178$) or PBMCs (Th17/CD4⁺CD25⁺CD127^{lo} Treg; $Z = 2.610, P = 0.456$; Th17/Foxp3⁺ Treg; $Z = 2.732, P = 0.435$) among the four groups (Figure 3 D, E, F, G).

3.5. Treg and Th17 cells and related molecules at the maternal-fetal interface in URSA

Flow cytometry was also used to detect the proportion of

Treg cells and Th17 cells and the expression of related molecules in decidual immune cells. Results indicated that the proportion of CD4⁺ T cells in decidual immune cells did not differ significantly between women with URSA and women with a normal pregnancy ($P > 0.05$) (Figure 4 A, B, C). The percentage of CD25⁺CD127^{lo} cells and Foxp3⁺ cells in CD4⁺ T cells was significantly lower in women with URSA than that in women in the early stage of normal pregnancy ($P < 0.05$) (Figure 4 D, E, F, G, H, I).

IL-10 expression by CD4⁺CD25⁺CD127^{lo} T cells did not differ significantly between women with URSA and women with a normal pregnancy ($P > 0.05$) (Figure 5 A, B, C). However, women with URSA had a significantly lower expression of CTLA-4 MFI on CD4⁺CD25⁺CD127^{lo} cells or CD4⁺Foxp3⁺ cells

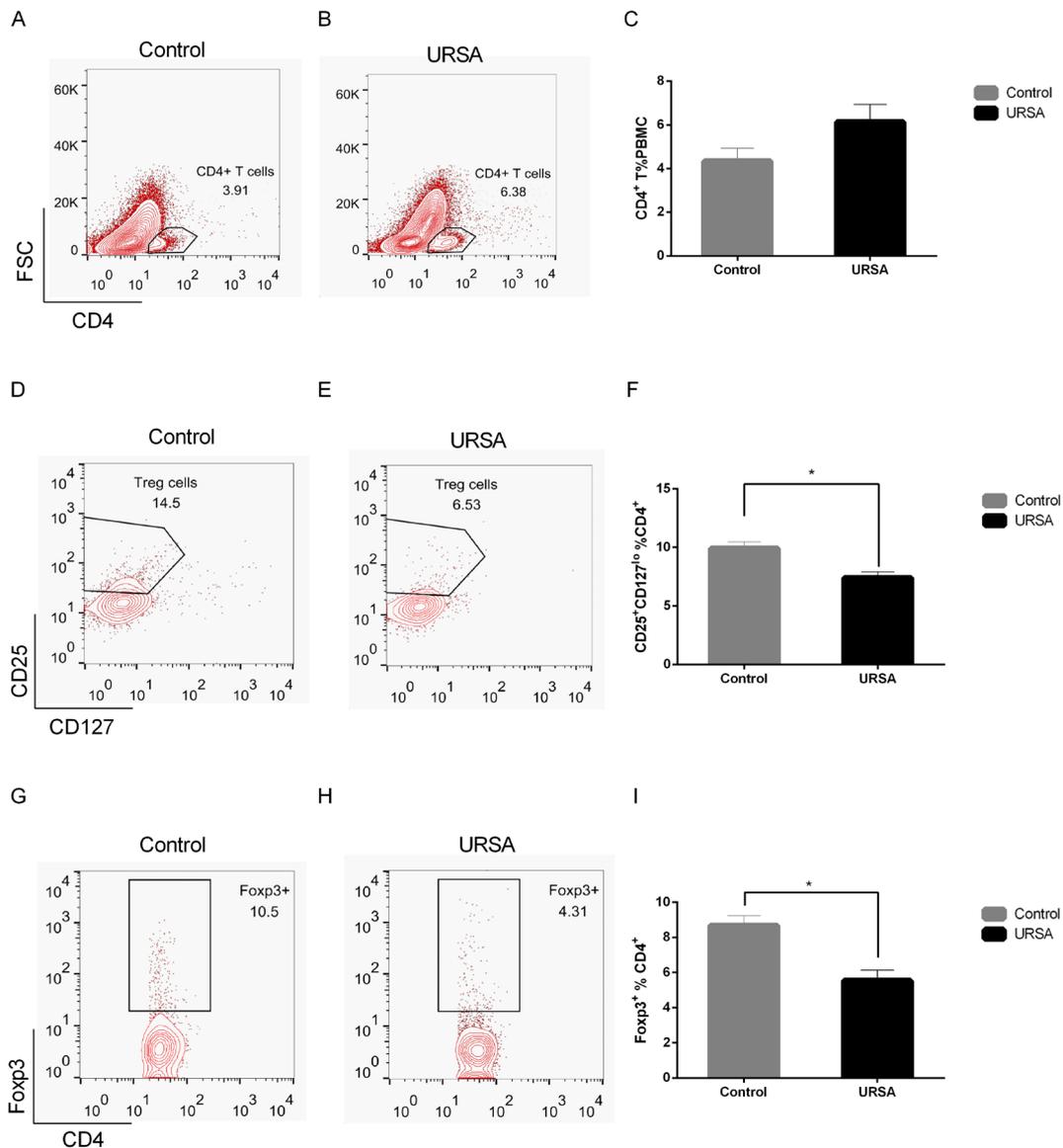


Figure 4. Proportion of CD4⁺ T cells and Treg cells in decidual immune cells. (A, B, C) There was no significant difference in the proportion of CD4⁺ T cells in the decidual immune cells between women with URSA and women with a normal pregnancy. (D, E, F) Women with URSA had a significantly lower percentage of CD25⁺CD127^{lo}/CD4⁺ T cells than their counterparts. (G, H, I) Women with URSA had a significantly lower percentage of Foxp3⁺/CD4⁺ T cells. (The FACS picture represents an individual sample. Data represent the Mean ± SEM. * $p < 0.05$.)

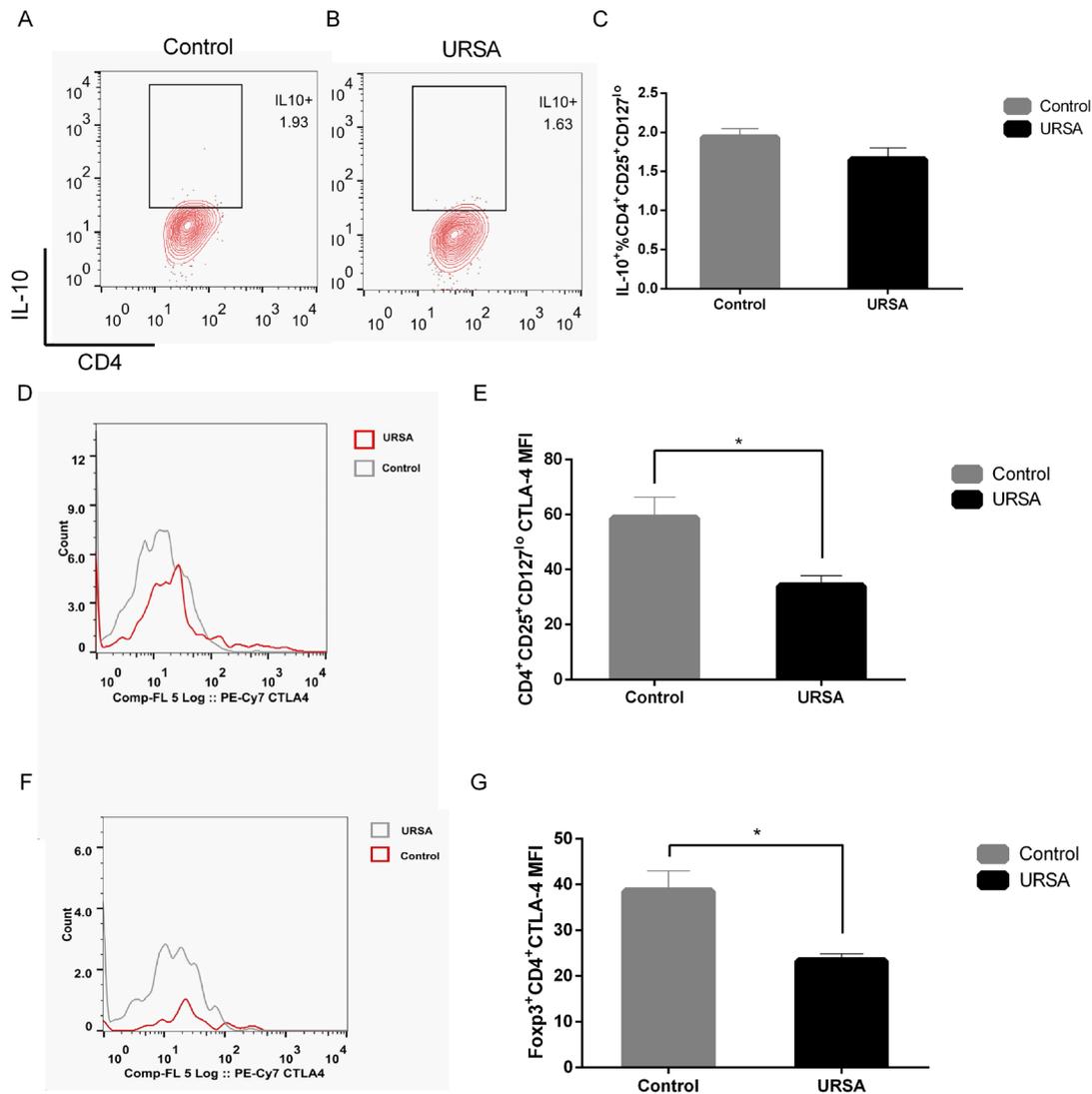


Figure 5. Expression of related molecules in decidual immune cells. (A, B, C) There was no significant difference in the IL-10 expression of CD4⁺CD25⁺CD127^{lo} T cells between women with URSA and women with a normal pregnancy. (D, E) Women with URSA had a significantly lower MFI of CD4⁺CD25⁺CD127^{lo} CTLA-4 expression than other groups. (F, G) Women with URSA had a significantly lower MFI of Foxp3⁺ CTLA-4 expression. (The FACS picture represents an individual sample. Data represent the Mean \pm SEM. * $p < 0.05$.)

compared to the control group respectively ($P < 0.05$) (Figure 5 D, E, F, G).

The percentage of IL-17⁺ cells among CD4⁺ T cells was slightly higher in women with URSA than that in women with a normal pregnancy, but the difference was not statistically significant ($P > 0.05$) (Figure 6 A, B, C). The ratio of Th17/CD25⁺CD127^{lo} Treg and Th17/Foxp3⁺ Treg in women with URSA was significantly higher than that in women with a normal pregnancy ($P < 0.05$) (Figure 6 D, E).

3.6. Levels of mRNA and protein of related molecules at the maternal-fetal interface in URSA

RT-PCR was used to detect the level of transcription of related molecules at the maternal-fetal interface in

women with URSA. Results indicated that levels of IL-6 and TNF- α mRNA in decidual tissue from women with URSA were significantly higher than those in women with a normal pregnancy ($P < 0.05$). Levels of Foxp3, CTLA-4, and TGF- β 1 mRNA were significantly lower ($P < 0.05$), while levels of IL-10, IL-17, IL-23, GITR, and ROR γ t mRNA did not statistically differ between the two groups ($P > 0.05$) (Figure 7).

The level of expression of related proteins in decidua was detected with ELISA. The levels of IL-6, IL-23, and TNF- α protein in women with URSA were significantly higher than those in women in the early stage of normal pregnancy ($P < 0.05$) and the level of TGF- β 1 protein in women with URSA was significantly lower ($P < 0.05$), while the levels of IL-17 and IL-10 did not statistically differ between the two groups ($P > 0.05$) (Figure 8).

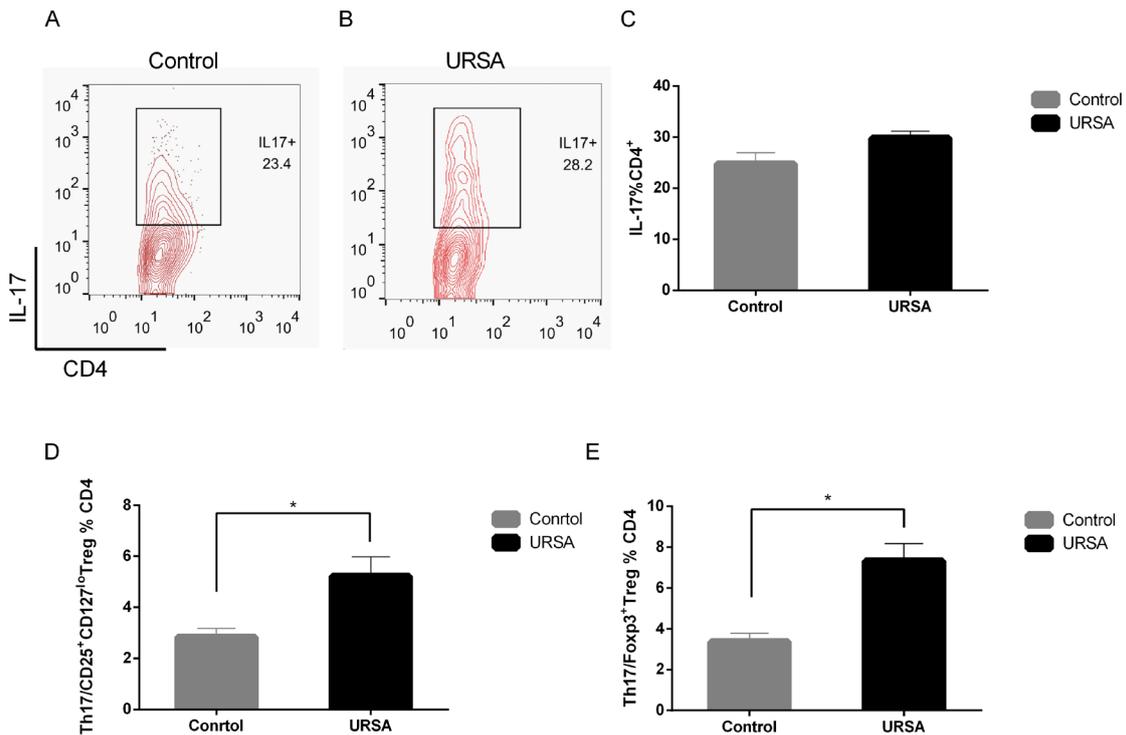


Figure 6. Proportion of Th17 cells and the ratio of Th17/Treg cells in decidual immune cells. (A, B, C) There was no significant difference in the percentage of IL-17⁺ cells in CD4⁺ T cells between women with URSA and women with a normal pregnancy. (D) Women with URSA had a significantly higher proportion of Th17/CD4⁺CD25⁺CD127^{lo} Treg cells than that in the control group. (E) Women with URSA had a significantly higher ratio of Th17/Foxp3⁺ Treg cells. (The FACS picture represents an individual sample. Data represent the Mean ± SEM. **p* < 0.05).

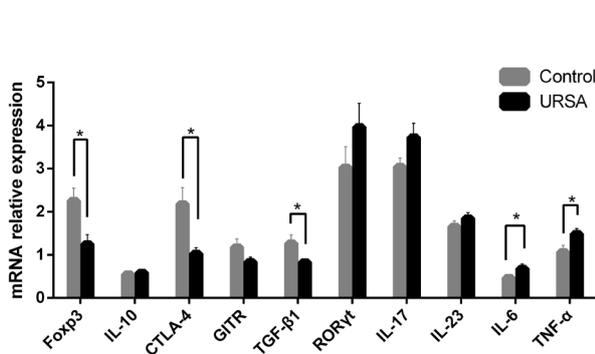


Figure 7. Levels of mRNA of related functional molecules at the maternal-fetal interface in URSA. (Data represent the Mean ± SEM. **p* < 0.05.)

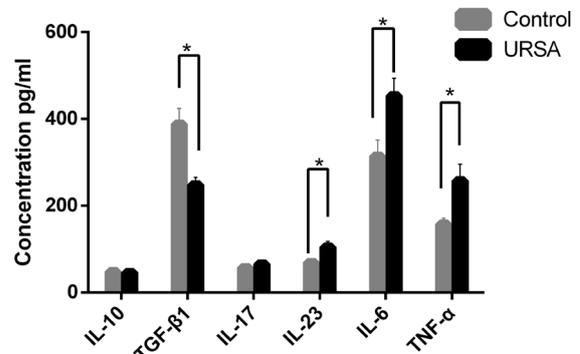


Figure 8. Levels of proteins of related functional molecules at the maternal-fetal interface in URSA. (Data represent the mean ± SEM. **p* < 0.05.)

4. Discussion

The aim of this study was to determine the pattern of Th17/Treg cells and related molecules in early pregnant women with a history of URSA. The immune system finely regulates the immune response in the human body to both fight infection and to protect itself from antigens (19). A successful pregnancy relies on the immune tolerance response, which enables the mother to tolerate the semi-allogeneic fetus. Without a delicate maternal-fetal immune balance, a series of pregnancy-related complications will occur, such as spontaneous abortion, preterm birth, and pre-eclampsia.

Several studies have indicated that a deficiency in Treg cells results in implantation failures and miscarriages in mice (20,21). The abortion rate can be reduced by transforming Treg cells in abortion-prone mice (22,23). In human pregnancy, Treg cells are found both in peripheral blood and deciduae. Somerset *et al.* (24) and Liu *et al.* (18) found that Treg cells in PBMCs significantly increased in the first trimester compared to the cell count in non-pregnant women. Most studies have noted an augmentation of Treg cells during the first and second trimester of pregnancy (25), while diminished levels of Treg cells have been reported in the second trimester of pregnancy (26). In the third trimester, Treg

levels decreased and further declined after labor (27). Studies have also indicated that the levels of Treg cells were significantly higher in normal pregnancy than those in spontaneous abortions (28,29). In assisted reproductive techniques (ART), the levels of Treg cells were higher in women who had a successful pregnancy and live birth than those in women with a failed pregnancy (30,31). The current study also found that peripheral Treg cells increased in PCs compared to NPCs. However, the elevation of Treg cells disappeared when pregnancy occurred in women with a history of URSA. Moreover, the proportion of Treg cells was significantly lower in women with URSA than that in women in the early stage of normal pregnancy both in the peripheral blood and at the maternal-fetal interface. This indicates that the immune response in pregnant women with a history of URSA differs from that in pregnant controls. The distinct pattern of immune response in women with URSA suggests that women with URSA are less likely to develop maternal-fetal immune tolerance during re-pregnancy.

IL-10, which is mainly secreted by Treg cells, is known to be an effective immune-regulating cytokine and an inhibitor of inflammatory cytokine synthesis. CTLA-4, a surface marker of Treg cells, may exhibit regulatory action in a contact-dependent manner. However, the function of IL-10 and CTLA-4 during pregnancy is still debated. Several studies found that IL-10 controlled the inflammatory process in pregnancy and that any change in the level of IL-10 might cause a disturbed pregnancy including URSA (32,33), while another study suggested that a deficiency in IL-10 had no effect on a disturbed pregnancy (34). Treg cells did not provide protection through CTLA-4 in a murine model of fetal tolerance (35,36). In contrast, a human study by Jin *et al.* proposed a role for CTLA-4 in Treg function (37). In the current study, there was no significant change in IL-10 expression either in peripheral blood or at the maternal-fetal interface after pregnancy, regardless of whether the woman had a history of URSA or not. Expression of CTLA-4 at the maternal-fetal interface decreased in women with URSA compared to that in pregnant controls, while peripheral CTLA-4 did not differ significantly among the groups. Further work is required to explore the mechanism of CTLA-4 in URSA.

IL-17, which is mainly secreted by Th17 cells, is a pro-inflammatory cytokine. The role of IL-17 and Th17 cells during pregnancy remains in question. Nakashima *et al.* found that the number of circulating Th17 cells did not change during different stages of pregnancy (non-pregnancy, first trimester, second trimester, or third trimester) (38). Liu *et al.* found that circulating Th17 cells did not differ markedly between women in the early stage of normal pregnancy and non-pregnant women (18). Another study reported that the proportion of Th17 cells in PBMCs was elevated in women in the early stage of normal pregnancy compared to that in

non-pregnant women (16). Lissauer *et al.* found that Th17 cell frequencies were remarkably heterogeneous among subjects, and a cross-sectional analysis revealed no difference in the percentage of Th17 cells in different stages of gestation (39). A longitudinal analysis depicted a 60% decline in Th17 cells between the first and second trimester of pregnancy. Since the uterine cavity is not sterile, the stable level of Th17 cells may also help to fight against extracellular pathogens. The abnormal elevation of IL-17 at the maternal-fetal interface might result in a miscarriage in mice, and a transfer of Treg cells before mating or administration of anti-IL-17 antibody might prevent URSA (40). The level of IL-17 is higher in peripheral blood in women with URSA compared to that in non-pregnant women (41). Treg cells decreased and Th17 cells increased during the window of implantation in women with a history of URSA (42,43). However, Nakashima *et al.* found that the proportion of IL-17 in a missed abortion did not statistically differ from that in normal pregnancy and that IL-17 was significantly higher in an inevitable abortion with active bleeding than that in a pregnancy with a normal decidua (44). Thus, IL-17 might not be responsible for an abortion in the initial stage, but it might be responsible in the later stage when inflammation develops. Thus, the exact role of Th17 cells and IL-17 in maternal-fetal tolerance is still uncertain. The current study found that IL-17⁺ CD4⁺ T cells were present in peripheral blood in both pregnant and non-pregnant women and that the proportion of those cells did not differ significantly between non-pregnant and pregnant women. At the maternal-fetal interface, there was no significant difference in the proportion of Th17 cells, and there was no significant difference in the level of IL-17 protein between women with URSA and women with a normal pregnancy.

An imbalance in Th17/Treg cells may be involved in autoimmune, inflammatory, and allergic conditions and pregnancy-related complications. The ratio of Th17/Treg cells was found to increase in women with URSA compared to that in healthy controls (45,46). The current results indicated that the ratio of peripheral Th17/Treg remained stable among the four groups while women with URSA had higher Th17/Treg ratio at the maternal-fetal interface than that in women with a normal pregnancy. This finding suggests that a balance in Th17 and Treg cells might be crucial to maternal-fetal immune tolerance.

5. Conclusion

The current study determined the pattern of immune cells and related molecules in URSA at the maternal-fetal interface as well as in peripheral blood. Results indicated that peripheral Treg cells did not increase when women with a history of URSA became pregnant, and this pattern was distinct from that in women with no history of URSA, suggesting that an alteration in

Treg cells might be involved in women with a history of URSA becoming pregnant. The decrease in Treg cells, CTLA-4, and TGF- β 1 as well as the increase in levels of IL-6 and TNF- α and the Th17/Treg ratio at the maternal-fetal interface might contribute to inappropriate maternal-fetal immune tolerance in URSA. Recognition of the immune cell pattern in women with a history of URSA is important to understanding the complicated mechanism of human reproduction. Two limitations of this study are the small sample size and lack of data on pregnancy outcomes. Further studies are required to elucidate the mechanism of immune tolerance.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (grant no. 31571196 to Ling Wang), the 2015 Program to Guide Medicine ("YixueYindao") of the Shanghai Municipal Science and Technology Commission grant no. 15401932200 to Ling Wang), the FY2008 JSPS Postdoctoral Fellowship for Foreign Researchers (P08471 to Ling Wang), the National Natural Science Foundation of China (grant no. 30801502 to Ling Wang), the Shanghai Pujiang Program (grant no. 11PJ1401900 to Ling Wang), the Shanghai Project for Development of Leading Disciplines-Integrative Medicine (grant no. 20150407) and the Science Foundation for Young Scientists of the Shanghai Municipal Commission of Health and Family Planning (grant no. 20174Y0224 to Na Zhang).

References

1. Sugiura-Ogasawara M, Ozaki Y and Suzumori N. Management of recurrent miscarriage. *J Obstet Gynaecol Res.* 2014; 40:1174-1179.
2. Li TC, Makris M, Tomsu M, Tuckerman E, Laird S. Recurrent miscarriage: Aetiology, management and prognosis. *Hum Reprod Update.* 2002; 8:463-481.
3. Saini V, Arora S, Yadav A, Bhattacharjee J. Cytokines in recurrent pregnancy loss. *Clin Chim Acta.* 2011; 412:702-708.
4. Saito S, Nakashima A, Shima T, Ito M. Th1/Th2/Th17 and regulatory T-cell paradigm in pregnancy. *Am J Reprod Immunol.* 2010; 63:601-610.
5. Wegmann TG, Lin H, Guilbert L, Mosmann TR. Bidirectional cytokine interactions in the maternal-fetal relationship: Is successful pregnancy a TH2 phenomenon? *Immunol Today.* 1993; 14:353-356.
6. Chaouat G, Ledee-Bataille N, Zourbas S, Ostojic S, Dubanchet S, Martal J, Frydman R. Cytokines, implantation and early abortion: Re-examining the Th1/Th2 paradigm leads to question the single pathway, single therapy concept. *Am J Reprod Immunol.* 2003; 50:177-186.
7. Sakaguchi S, Sakaguchi N, Asano M, Itoh M, Toda M. Immunologic self-tolerance maintained by activated T cells expressing IL-2 receptor α -chains (CD25). Breakdown of a single mechanism of self-tolerance causes various autoimmune diseases. *J Immunol.* 1995; 155:1151-1164.
8. Liu W, Putnam AL, Xu-Yu Z, Szot GL, Lee MR, Zhu S, Gottlieb PA, Kapranov P, Gingeras TR, Fazekas DSGB, Clayberger C, Soper DM, Ziegler SF, Bluestone JA. CD127 expression inversely correlates with FoxP3 and suppressive function of human CD4⁺ T reg cells. *J Exp Med.* 2006; 203:1701-1711.
9. Takahashi T, Tagami T, Yamazaki S, Uede T, Shimizu J, Sakaguchi N, Mak TW, Sakaguchi S. Immunologic self-tolerance maintained by CD25⁺CD4⁺ regulatory T cells constitutively expressing cytotoxic T lymphocyte-associated antigen 4. *J Exp Med.* 2000; 192:303-310.
10. McHugh RS and Shevach EM. Cutting edge: depletion of CD4⁺CD25⁺ regulatory T cells is necessary, but not sufficient, for induction of organ-specific autoimmune disease. *J Immunol.* 2002; 168:5979-5983.
11. Nie H, Zheng Y, Li R, Guo TB, He D, Fang L, Liu X, Xiao L, Chen X, Wan B, Chin YE, Zhang JZ. Phosphorylation of FOXP3 controls regulatory T cell function and is inhibited by TNF- α in rheumatoid arthritis. *Nat Med.* 2013; 19:322-328.
12. Corthay A. How do regulatory T cells work? *Scand J Immunol.* 2009; 70:326-336.
13. Manni ML, Robinson KM and Alcorn JF. A tale of two cytokines: IL-17 and IL-22 in asthma and infection. *Expert Rev Respir Med.* 2014; 8:25-42.
14. Wichner K, Stauss D, Kampfrath B, Kruger K, Muller G, Rehm A, Lipp M, Hopken UE. Dysregulated development of IL-17- and IL-21-expressing follicular helper T cells and increased germinal center formation in the absence of ROR γ mat. *FASEB J.* 2016; 30:761-774.
15. Schumacher A and Zenclussen AC. Regulatory T cells: Regulators of life. *Am J Reprod Immunol.* 2014; 72:158-170.
16. Wu HX, Jin LP, Xu B, Liang SS, Li DJ. Decidual stromal cells recruit Th17 cells into decidua to promote proliferation and invasion of human trophoblast cells by secreting IL-17. *Cell Mol Immunol.* 2014; 11:253-262.
17. Xu WM, Xiao ZN, Wang XB, Huang Y. IL-17 induces fetal loss in a CBA/JxBALB/c mouse model, and an anti-IL-17 antibody prevents fetal loss in a CBA/JxDBA/2 mouse model. *Am J Reprod Immunol.* 2016; 75:51-58.
18. Liu YS, Wu L, Tong XH, Wu LM, He GP, Zhou GX, Luo LH, Luan HB. Study on the relationship between Th17 cells and unexplained recurrent spontaneous abortion. *Am J Reprod Immunol.* 2011; 65:503-511.
19. Chatterjee P, Chiasson VL, Bounds KR, Mitchell BM. Regulation of the anti-inflammatory cytokines interleukin-4 and interleukin-10 during pregnancy. *Front Immunol.* 2014; 5:253.
20. Shima T, Sasaki Y, Itoh M, Nakashima A, Ishii N, Sugamura K, Saito S. Regulatory T cells are necessary for implantation and maintenance of early pregnancy but not late pregnancy in allogeneic mice. *J Reprod Immunol.* 2010; 85:121-129.
21. Rowe JH, Ertelt JM, Aguilera MN, Farrar MA, Way SS. Foxp3⁺ regulatory T cell expansion required for sustaining pregnancy compromises host defense against prenatal bacterial pathogens. *Cell Host Microbe.* 2011; 10:54-64.
22. Yin Y, Han X, Shi Q, Zhao Y, He Y. Adoptive transfer of CD4⁺CD25⁺ regulatory T cells for prevention and treatment of spontaneous abortion. *Eur J Obstet Gynecol Reprod Biol.* 2012; 161:177-181.

23. Zenclussen AC, Gerlof K, Zenclussen ML, Ritschel S, Zambon BA, Fest S, Hontsu S, Ueha S, Matsushima K, Leber J, Volk HD. Regulatory T cells induce a privileged tolerant microenvironment at the fetal-maternal interface. *Eur J Immunol.* 2006; 36:82-94.
24. Somerset DA, Zheng Y, Kilby MD, Sansom DM, Drayson MT. Normal human pregnancy is associated with an elevation in the immune suppressive CD25+ CD4+ regulatory T-cell subset. *Immunology.* 2004; 112:38-43.
25. Xiong YH, Yuan Z and He L. Effects of estrogen on CD4⁺CD25⁺ regulatory T cell in peripheral blood during pregnancy. *Asian Pac J Trop Med.* 2013; 6:748-752.
26. Mjosberg J, Svensson J, Johansson E, Hellstrom L, Casas R, Jenmalm MC, Boij R, Matthiesen L, Jonsson JI, Berg G, Ernerudh J. Systemic reduction of functionally suppressive CD4^{dim}CD25^{high}Foxp3⁺ Tregs in human second trimester pregnancy is induced by progesterone and 17beta-estradiol. *J Immunol.* 2009; 183:759-769.
27. Seol HJ, Oh MJ, Lim JE, Jung NH, Yoon SY, Kim HJ. The role of CD4⁺CD25^{bright} regulatory T cells in the maintenance of pregnancy, premature rupture of membranes, and labor. *Yonsei Med J.* 2008; 49:366-371.
28. Inada K, Shima T, Nakashima A, Aoki K, Ito M, Saito S. Characterization of regulatory T cells in decidua of miscarriage cases with abnormal or normal fetal chromosomal content. *J Reprod Immunol.* 2013; 97:104-111.
29. Mei S, Tan J, Chen H, Chen Y, Zhang J. Changes of CD4⁺CD25^{high} regulatory T cells and FOXP3 expression in unexplained recurrent spontaneous abortion patients. *Fertil Steril.* 2010; 94:2244-2247.
30. Lu Y, Zhang F, Zhang Y, Zeng B, Hu L, Liao A. Quantitative reduction of peripheral CD4⁺ CD25⁺ FOXP3⁺ regulatory T cells in reproductive failure after artificial insemination by donor sperm. *Am J Reprod Immunol.* 2013; 69:188-193.
31. Zhou J, Wang Z, Zhao X, Wang J, Sun H, Hu Y. An increase of Treg cells in the peripheral blood is associated with a better in vitro fertilization treatment outcome. *Am J Reprod Immunol.* 2012; 68:100-106.
32. Zhu L, Chen H, Liu M, Yuan Y, Wang Z, Chen Y, Wei J, Su F, Zhang J. Treg/Th17 cell imbalance and IL-6 profile in patients with unexplained recurrent spontaneous abortion. *Reprod Sci.* 2017; 24:882-890.
33. Choi YK and Kwak-Kim J. Cytokine gene polymorphisms in recurrent spontaneous abortions: A comprehensive review. *Am J Reprod Immunol.* 2008; 60:91-110.
34. Svensson L, Arvola M, Sallstrom MA, Holmdahl R, Mattsson R. The Th2 cytokines IL-4 and IL-10 are not crucial for the completion of allogeneic pregnancy in mice. *J Reprod Immunol.* 2001; 51:3-7.
35. Schumacher A, Wafula PO, Bertoja AZ, Sollwedel A, Thuere C, Wollenberg I, Yagita H, Volk HD, Zenclussen AC. Mechanisms of action of regulatory T cells specific for paternal antigens during pregnancy. *Obstet Gynecol.* 2007; 110:1137-1145.
36. Schumacher A, Wafula PO, Teles A, El-Mousleh T, Linzke N, Zenclussen ML, Langwisch S, Heinze K, Wollenberg I, Casalis PA, Volk HD, Fest S, Zenclussen AC. Blockage of heme oxygenase-1 abrogates the protective effect of regulatory T cells on murine pregnancy and promotes the maturation of dendritic cells. *PLoS One.* 2012; 7:e42301.
37. Jin LP, Chen QY, Zhang T, Guo PF, Li DJ. The CD4⁺CD25⁺ bright regulatory T cells and CTLA-4 expression in peripheral and decidual lymphocytes are down-regulated in human miscarriage. *Clin Immunol.* 2009; 133:402-410.
38. Nakashima A, Ito M, Yoneda S, Shiozaki A, Hidaka T, Saito S. Circulating and decidual Th17 cell levels in healthy pregnancy. *Am J Reprod Immunol.* 2010; 63:104-109.
39. Lissauer D, Goodyear O, Khanum R, Moss PA, Kilby MD. Profile of maternal CD4 T-cell effector function during normal pregnancy and in women with a history of recurrent miscarriage. *Clin Sci (Lond).* 2014; 126:347-354.
40. Wang WJ, Liu FJ, Xin-Liu, Hao CF, Bao HC, Qu QL, Liu XM. Adoptive transfer of pregnancy-induced CD4⁺CD25⁺ regulatory T cells reverses the increase in abortion rate caused by interleukin 17 in the CBA/JxBALB/c mouse model. *Hum Reprod.* 2014; 29:946-952.
41. Lee SK, Kim JY, Hur SE, Kim CJ, Na BJ, Lee M, Gilman-Sachs A, Kwak-Kim J. An imbalance in interleukin-17-producing T and Foxp3⁺ regulatory T cells in women with idiopathic recurrent pregnancy loss. *Hum Reprod.* 2011; 26:2964-2971.
42. Saifi B, Rezaee SA, Tajik N, Ahmadpour ME, Ashrafi M, Vakili R, SoleimaniAsl S, Aflatoonian R, Mehdizadeh M. Th17 cells and related cytokines in unexplained recurrent spontaneous miscarriage at the implantation window. *Reprod Biomed Online.* 2014; 29:481-489.
43. Sereshki N, Gharagozloo M, Ostadi V, Ghahiri A, Roghaei MA, Mehrabian F, Andalib AA, Hassanzadeh A, Hosseini H, Rezaei A. Variations in T-helper 17 and regulatory T cells during the menstrual cycle in peripheral blood of women with recurrent spontaneous abortion. *Int J Fertil Steril.* 2014; 8:59-66.
44. Nakashima A, Ito M, Shima T, Bac ND, Hidaka T, Saito S. Accumulation of IL-17-positive cells in decidua of inevitable abortion cases. *Am J Reprod Immunol.* 2010; 64:4-11.
45. Wu L, Luo LH, Zhang YX, Li Q, Xu B, Zhou GX, Luan HB, Liu YS. Alteration of Th17 and Treg cells in patients with unexplained recurrent spontaneous abortion before and after lymphocyte immunization therapy. *Reprod Biol Endocrinol.* 2014; 12:74.
46. Lee SK, Kim JY, Lee M, Gilman-Sachs A, Kwak-Kim J. Th17 and regulatory T cells in women with recurrent pregnancy loss. *Am J Reprod Immunol.* 2012; 67:311-318.

(Received January 30, 2018; Revised March 20, 2018; Accepted March 22, 2018)

Renal protective effect of Paeoniflorin by inhibition of JAK2/STAT3 signaling pathway in diabetic mice

Xinyu Li[§], Yan Wang[§], Kun Wang, Yonggui Wu*

Department of Nephrology, the First Affiliated Hospital, Anhui Medical University, Hefei, Anhui, China.

Summary

Paeoniflorin is the main bioactive components of the root of *P. lactiflora* Pall., and has been widely used as an anti-inflammation and immunomodulatory agent. However, the effect and mechanisms of Paeoniflorin in diabetic nephropathy (DN) remains to be elucidated. In the present study, streptozotocin (STZ)-induced type 1 diabetic mice model was used to investigate the protective effect of Paeoniflorin and the role of the Janus kinase (JAK) 2/ signal transducer (STAT) 3 signaling pathway on DN. After treatment with Paeoniflorin at a dose of 25, 50 and 100 mg/kg once a day for 12 weeks, both the functional and histological damage to diabetic mice kidney had been attenuated significantly. Additionally, these reno-protective effects were associated with alleviating macrophage infiltration and inflammatory factors expression as well as suppression of the JAK2/STAT3 signaling pathway. These data reveal that Paeoniflorin attenuates renal lesions in diabetic mice and these protective effects may be associated with the prevention of macrophage infiltration and inhibition of the JAK2/STAT3 signaling pathway.

Keywords: Paeoniflorin, diabetic nephropathy, JAK2/STAT3, inflammation, macrophage

1. Introduction

As a major severe complication of diabetes mellitus (DM), diabetic nephropathy (DN) is the leading cause of end-stage renal disease (ESRD) and imposes high social and economic burden worldwide (1). A study from Zhang *et al.* (2) indicated that with the increasing prevalence of DM, chronic kidney disease related to diabetes had become more common than chronic kidney disease related to glomerulonephritis in China. Increasing evidence showed that the pathogenesis of DN was associated with the interactions between metabolic and hemodynamic alteration, oxidative stress, inflammation, activation of the renin-angiotensin-aldosterone system (RAAS) and other factors (3). Studies have indicated that chronic low

grade inflammatory reactions play critical roles in the initiation and progression of DN (4). Presently, available therapies such as intensive hyperglycemia, blood pressure control, and inhibition of the RAAS have been shown merely to slow but insufficiently prevent the progression of DN (5). Thus, it is of pressing need to identify novel therapeutic targets that might lead to the prevention of DN.

The Janus kinase (JAK)/signal transducer and activator of transcription (STAT) signaling pathway regulates a broad range of biological effects such as cell proliferation, differentiation, inflammation, and apoptosis (6). The JAK/STAT pathway transmits extracellular ligand signals directly to the nucleus to induce a variety of cellular responses. Currently, there are four members of the JAK family (JAK1, JAK2, JAK3, and Tyrosine Kinase 2) and seven STATs (STAT1, 2, 3, 4, 5a, 5b and 6) identified in mammals (7). Recent studies have shown that the JAK/STAT pathway is activated in numerous acute and chronic renal diseases, moreover, JAK2 and STAT3 are the JAK and STAT forms that have been most clearly identified in the onset and progression of DN (8). The accumulated evidence for JAK2/STAT3 activation in the pathogenesis of DN implied a novel therapeutic target for potential

Released online in J-STAGE as advance publication April 19, 2018.

[§]These authors contributed equally to this work.

*Address correspondence to:

Dr. Yonggui Wu, Department of Nephropathy, the First Affiliated Hospital, Anhui Medical University, No.218 Jixi road, Shushan District, Hefei, Anhui 230022, China.
E-mail: wuyonggui@medmail.com.cn

intervention in this disease.

Paeoniflorin is a monoterpene glycoside extracted from the dried root of *P. lactiflora* Pall. As the principle bioactive component (> 90 %) of the Chinese traditional medicine total glucosides of paeony (TGP) (9), Paeoniflorin exerts numerous pharmacological effects including anti inflammation, immunomodulation, neuroprotective and antitumor effects (10-13). Fu *et al.* (14) reported that Paeoniflorin prevented DN by inhibiting activation of nuclear factor- κ B (NF- κ B) and renal macrophage infiltration in streptozotocin (STZ)-induced diabetic model rats. Furthermore, our previous studies demonstrated that Paeoniflorin inhibited toll-like receptor (TLR) 2/4-mediated inflammation and exerted a renoprotective role on type 2 diabetic nephropathy in both *in vitro* and *in vivo* studies (15). However, the exact mechanisms of Paeoniflorin in DN still remains unclear. Additionally, whether the renoprotective effects of Paeoniflorin in DN are induced *via* the JAK2/STAT3 signaling pathway requires further exploration.

Therefore, the aim of the present study was to evaluate whether Paeoniflorin could alleviate renal damage in STZ-induced type 1 diabetes mellitus using C57BL/6J mice. Moreover, we hypothesized that Paeoniflorin might exert renoprotective effects by inhibiting the JAK2/STAT3 signaling pathway.

2. Materials and Methods

2.1. Drugs and reagents

Paeoniflorin [(C₂₃H₂₈O₁₁, MW: 480.45, purity = 98.78% (high-performance liquid chromatography), lethal dose = 9,530 mg/kg, Figure 1)] was purchased from Nanjing Goren Biotechnology Co., Ltd (Nanjing, China). STZ was purchased from Sigma-Aldrich Co. (St Louis, MO, USA). The rabbit anti-JAK2, STAT3, phosphorylated JAK2(p-JAK2), phosphorylated STAT3(p-STAT3) antibodies were bought from Cell Signaling Technology (Beverly, MA, USA) and CD68 was bought from Santa Cruz Biotechnology (Santa Cruz, CA, USA). Anti- β -actin antibodies, anti-mouse IgG and anti-rabbit IgG horseradish-peroxidase (HRP) were brought from Wuhan Sanying Biotechnology (Wuhan, China). GAPDH and tumor necrotic factor- α (TNF- α) primers were obtained from the Shanghai Sangon Biotech Co., Ltd (Shanghai, China). Primers for inducible nitric oxide synthase (iNOS) (MQP029793), interleukin-1 β (IL-1 β) (MQP027422), and monocyte chemoattractant protein-1 (MCP-1) (MQP027672) were purchased commercially from GeneCopoeia, Inc. (Rockville, MD, USA).

2.2. Animals and experimental design

Male C57BL/6J mice ($n = 60$; 8-10 weeks old; weight between 18-20 g) were purchased from the Model

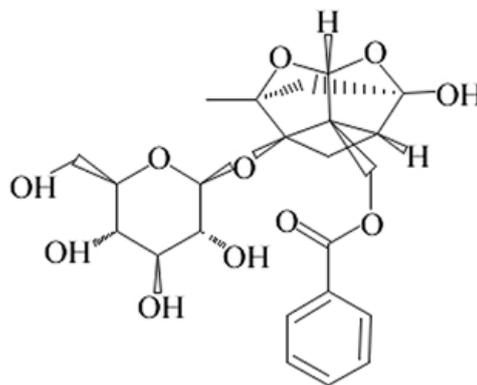


Figure 1. The chemical structure of Paeoniflorin.

Animal Research Centre of Nanjing University. All mice were housed under standard raising condition with room temperature (20-24°C), humidity (50-55%) with a 12/12 h light/dark cycle and free access to food and water in the Experimental Animal Center of Anhui Medical University.

After 1 week for acclimation, mice were randomly divided into five groups ($n = 12$ per group): (1) normal control group (Con group); (2) diabetic model group (DM group); (3) Paeoniflorin 25mg/kg treatment group (DM+PF 25 mg/kg); (4) Paeoniflorin 50 mg/kg treatment group (DM+PF 50mg/kg); (5) Paeoniflorin 100 mg/kg treatment group (DM+PF 100 mg/kg). The type 1 diabetic models were produced by intraperitoneal injection of STZ (50 mg/kg in 0.1 M citrate buffer, pH 4.5) daily for 5 consecutive days (16). Mice in the control group received an equal volume of the citrate buffer. Blood glucose was measured 1 week after STZ injection, and mice with blood glucose over 16.7 mmol/L were considered diabetic and selected for the following study. Paeoniflorin gavage was initiated once a day at a dose of 25, 50, and 100 mg/kg for 12 weeks while the Con group and the DM group received an equivalent volume of distilled water (17). After 12 weeks, blood and 24 h urine samples were collected. After that, all the mice were sacrificed and body weights were recorded. Kidney samples of mice were dissected, weighed and processed for further analyses. All animal investigation was approved by the Ethical Committee of Animal Research of Anhui Medical University, and the mice were sacrificed according to the Guide for the Care and Use of Laboratory Animals recommendations.

2.3. Measurement of general and metabolic parameters

Body weight (BW), kidney weight (KW) and blood glucose was recorded by an electronic scale and glucose analyzer. All Mice were housed individually in metabolic cages for 24 h to collect urinary samples before being sacrificed. The concentration of 24 h

urine albumin was determined by using a mouse microalbumin ELISA kit (Abcam Biotechnology, Cambridge, UK).

2.4. Histological examination

For histological and immunohistochemistry analysis, 4% paraformaldehyde-fixed kidney tissues were dehydrated through a graded series of ethanol, routine paraffin embedding, and then cut into 3 μ m sections to make tissue slides. Periodic acid-schiff (PAS) staining was performed under a light microscope. PAS staining was used for histological grading and the glomerular mesangial expansion index and tubulointerstitial damage index as previously described was measured randomly in 10 visual fields (18).

2.5. Immunohistochemistry (IHC) analysis

After being deparaffinized with xylene and hydrated with graded alcohol, the kidney tissue slides were placed in 3% hydrogen peroxide to block endogenous peroxidase activity. For antigen retrieval, kidney slides were transferred to 10 mmol/L citrate buffer solution (pH 6.0) at 100°C for 15 min. Then the slides were incubated at 37°C for 30 min with goat serum to prevent nonspecific binding and incubated with primary antibodies: anti-CD68 (1:50), anti-p-JAK2 (1:100), and anti-p-STAT3 (1:100) at the given dilutions overnight at 4°C. After washing, the slides were incubated with polyperoxidase-anti-mouse/rabbit IgG for 30 min at 37°C followed by color development with 3,3-diaminobenzidine (Sigma-Aldrich, St. Louis, MO, USA) and hematoxylin staining was used. The numbers of CD68-positive cells and the percentage of the p-JAK2 and the p-STAT3 brown positive staining areas in glomeruli and tubulointerstitium areas were calculated respectively in 10 random high-power (\times 400) fields. Quantitative analysis was performed using Image-Pro Plus 6.0 (Media Cybernetics, Rockville, MD, USA).

2.6. Western blot analysis

Protein samples used for Western blot analysis were extracted from mice renal samples and the protein concentration was determined with a Bio-Rad protein assay kit (Beyotime Institute of Biotechnology, Jiangsu, China). Equal amounts of proteins were separated by 8-12% sodium dodecylsulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to nitrocellulose membranes (Millipore, Billerica, MA, USA). The membranes were blocked with 5% nonfat milk in Tris-buffered saline and Tween 20 (TBST) buffer for 2 h and then incubated overnight at 4°C with different primary polyclonal antibodies against p-JAK2 (1:1,000), JAK2 (1:1,000), p-STAT3 (1:1,000), and STAT3 (1:1,000). After washing, the membranes

were incubated with HRP-labeled secondary antibody (1:35,000) at room temperature for 1 h. The blots were revealed using an enhanced chemiluminescence (ECL) kit (Amersham Life Science, Little Chalfont, UK) and the protein content was quantified and analyzed using a Leica Q500IW image analysis system (Leica Ltd., Cambridge, UK). β -actin served as a loading control. The band density was measured using Image J software.

2.7. Real-time PCR (RT-PCR)

Total RNA was extracted from kidney tissues using TRIzol reagent (Invitrogen, California, USA). The transcription of cDNA was synthesized with a standard reverse transcription reaction kit (Promega, Madison, WI, USA) and RT-PCR was performed using the SYBR Green PCR Master Mix kit (Bio-Rad Laboratories, Hercules, CA, USA). The amplification conditions were set as follows: initial hold steps (denaturation at 95°C for 10min) and 35 cycles of a 2-step PCR (95°C for 15s and 60°C for 30s). The primers used in RT-PCR were GAPDH (forward, 5'-GGTGAAGGTCGGTGTGAACG-3'; reverse, 5'-CTCGCTCCTGGAAGATGGTG-3'), TNF- α (forward, 5'-GCTGAGCTCAAACCCTGGTA-3'; reverse, 5'-CGGACTCCG CAAAGTCTAAG-3'), iNOS (MQP029793), IL-1 β (MQP027422), and MCP-1 (MQP027672). Primers were purchased from GeneCopoeia, Inc. The relative expression levels of the target PCR product was calculated after adjusting for GAPDH by using the 2- $\Delta\Delta$ Ct method (19).

2.8. Statistical analysis

SPSS software version 16.0 (SPSS Inc., Chicago, IL, USA) was used for statistical analysis. Data were expressed as mean \pm standard deviation (SD) and the results were analyzed by one-way analysis of variance (ANOVA). The difference between groups was tested by least significant difference (LSD) and the Levene method was used for homogeneity test of variance. A value of $p < 0.05$ was considered statistically significant.

3. Results

3.1. Paeoniflorin attenuates renal damage in diabetic kidneys

Initially, to investigate the effects of Paeoniflorin on diabetic kidney, we evaluated renal function by measuring general and metabolic parameters including blood glucose, KW/BW ratio and 24 h urinary albumin excretion rate (UAER). The DM group demonstrated elevated blood glucose, increased KW/BW ratio and the differences were significant compared with that of the control group. Paeoniflorin treatment for 12 weeks had

Table 1. General and metabolic parameters from five groups of mice models

Parameter	Con	DM	DM+PF 25 mg/ kg	DM+PF 50 mg/ kg	DM+PF 100 mg/ kg
Blood glucose (mmol/L)	5.78 ± 3.16	30.46 ± 2.45**	26.52 ± 3.62	26.33 ± 5.64	30.84 ± 5.07
KW/BW ratio (g/100g BW)	0.70 ± 0.07	0.84 ± 0.11*	0.79 ± 0.11 [#]	0.77 ± 0.05 [#]	0.77 ± 0.13 [#]
UAER (mg/24 h)	0.09 ± 0.02	1.69 ± 0.13**	1.25 ± 0.16 ^{##}	0.84 ± 0.15 ^{###}	0.32 ± 0.09 ^{###}

All the values are represented as mean ± SD ($n = 12$ per group). * $P < 0.05$, ** $P < 0.01$ vs. Con; [#] $P < 0.05$, ^{##} $P < 0.01$ vs. DM. Con, normal control group; DM, diabetic model group; PF, Paeoniflorin; KW/BW, kidney weight/body weight; UAER, urinary albumin excretion rate.

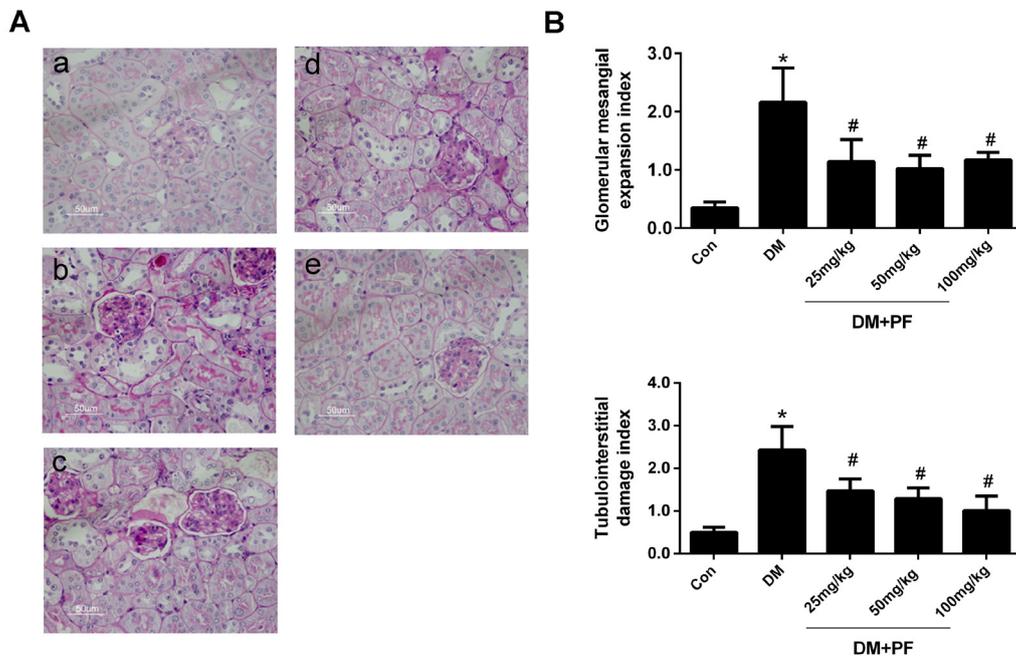


Figure 2. Histology features of renal tissues. (A) Representative images of kidney samples stained with PAS in the (A-a) Con, (A-b) DM, (A-c) DM+PF 25 mg/kg, (A-d) DM + PF 50 mg/kg and (A-e) DM + PF 100mg/kg groups. (B) Quantification of PAS staining. All the values are represented as mean ± SD of at least three independent experiments ($n = 12$ per group). * $p < 0.05$ vs. Con; [#] $p < 0.05$ vs. DM. Con, normal control group; DM, diabetic model group; PF, Paeoniflorin; PAS, periodic acid-schiff. Original magnification, $\times 400$ in A.

no significant effect on the hyperglycemia in diabetic mice. However, the KW/BW ratio of the DM group was significantly reduced after Paeoniflorin treatment in diabetic mice. Similarly, the diabetic mice had a remarkably higher 24 h UAER than those in the control group. After Paeoniflorin treatment for 12 weeks, the raised 24 h UAER of the diabetic group was attenuated significantly in a dose-dependent manner although their levels were still higher than those of the control group. These results showed that Paeoniflorin could preserve STZ-caused renal lesions (Table 1).

In addition to clinical renal function results, we further evaluated renal histopathological alterations in kidney samples. Results from PAS staining indicated that both of the glomerular mesangial expansion index and tubulointerstitial damage index of the DM group were particularly higher when compared with the control group (Figure 2B). By contrast, Paeoniflorin treatment for 12 weeks decreased pathological lesions index remarkably compared to the DM group (Figure 2A and B). Thus histological assessment further confirmed that Paeoniflorin treatment could ameliorate

progression of DN.

3.2. Paeoniflorin suppresses macrophage infiltration and macrophage-mediated inflammatory factors in diabetic kidneys

The induction of diabetes was associated with the recruitment, retention, and activation of macrophages in mice kidneys, as evidenced by increased expression of macrophage markers and proinflammatory factors. In this study, CD68 as a macrophage surface marker is used to investigate the role of Paeoniflorin in inflammatory responses related to DN in renal tissues. No obvious infiltration of CD68-positive macrophages was detected in the glomerular and tubulointerstitial areas of the normal control group. In contrast, we observed abnormal macrophage infiltration in the DM group. Paeoniflorin treatment markedly reduced such abnormal macrophage infiltration in the diabetic kidneys in a dose-dependent manner (Figure 3A and 3B).

To further evaluate whether Paeoniflorin modulates

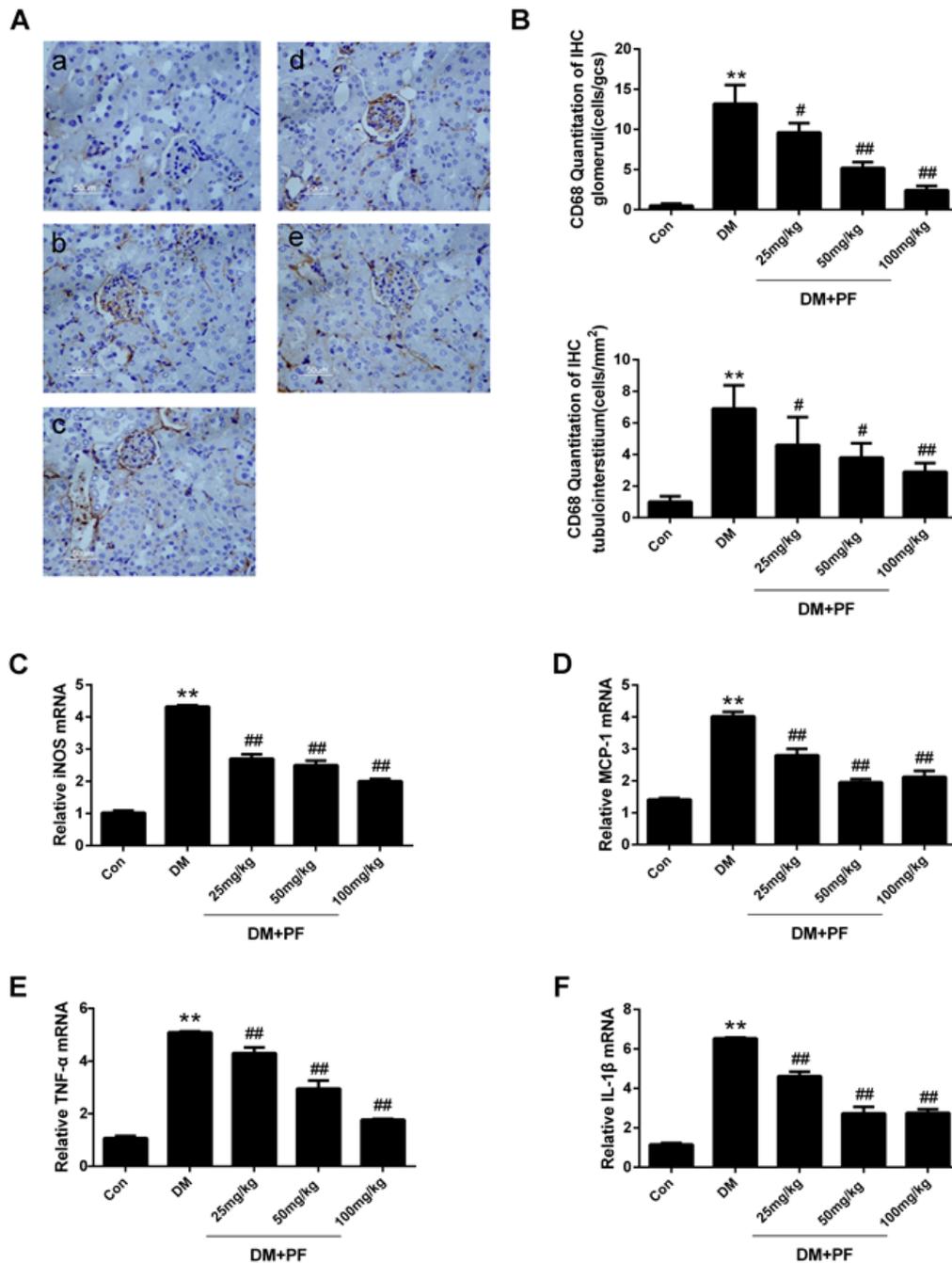


Figure 3. Paeniflorin inhibits STZ-induced macrophage infiltration and inflammatory cytokine expression in renal tissues. (A) Representative images of CD68 IHC staining of kidney samples in the (A-a) Con, (A-b) DM, (A-c) DM+PF 25 mg/kg, (A-d) DM+PF 50 mg/kg and (A-e) DM+PF 100 mg/kg groups. (B) Quantitative analyses of CD68 IHC staining. (C-F) mRNA expression of iNOS, MCP-1, TNF- α and IL1- β in renal tissues, as determined by real-time PCR. All values are represented as mean \pm SD of at least three independent experiments ($n = 12$ per group). * $p < 0.05$, ** $p < 0.01$ vs. Con; # $p < 0.05$, ## $p < 0.01$ vs. DM. STZ, streptozocin; Con, normal control group; DM, diabetic model group; PF, Paeniflorin; IHC, immunohistochemical; iNOS, inducible nitric oxide synthase; MCP-1, monocyte chemoattractant protein-1; TNF- α , tumor necrotic factor- α ; IL1- β , interleukin-1 β . Original magnification, $\times 400$ in A.

the functional stage of kidney macrophages, expression levels of iNOS were examined to distinguish between pro-inflammatory M1 and anti-inflammatory M2 phenotypes. The mRNA expression level of iNOS in the DM group increased remarkably compared to the control group while Paeniflorin treatment significantly repressed this overexpression (Figure 3C). Moreover,

we also tested the expression levels of proinflammatory cytokines including MCP-1, TNF- α and IL-1 β . Similarly, results indicated that Paeniflorin treatment inhibited the up-regulation of all of these cytokines in the STZ-induced vehicle group to a different extent (Figure 3D-F). These present results indicated that Paeniflorin could mitigate renal inflammation in

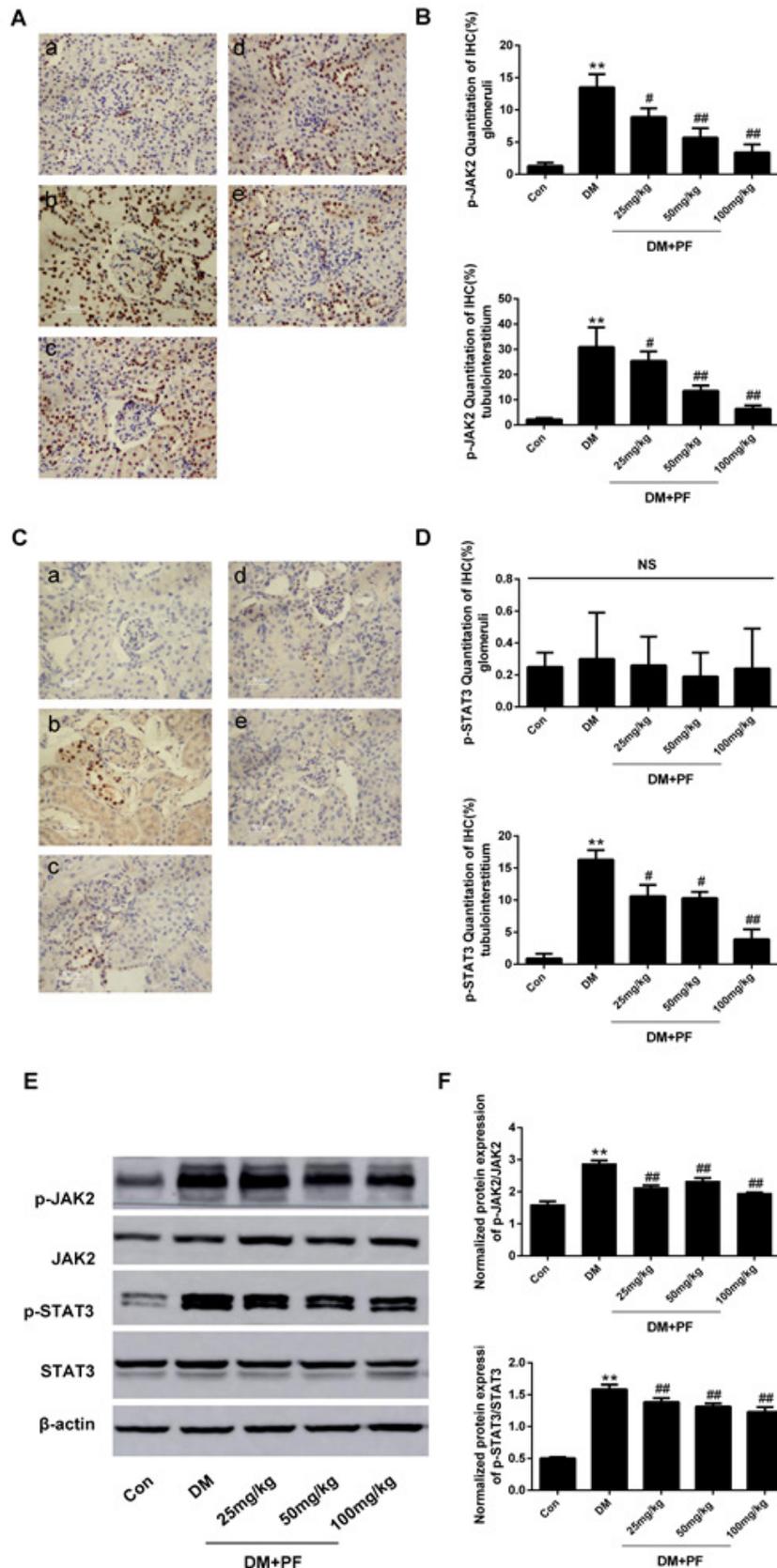


Figure 4. Effects of Paconiflorin on the expression of p-JAK2, p-STAT3, JAK2 and STAT3 in renal tissues. (A-B) Representative images and quantitative analyses of IHC staining for p-JAK2 expression in the (A-a) Con, (A-b) DM, (A-c) DM+PF 25 mg/kg, (A-d) DM+PF 50 mg/kg and (A-e) DM+PF 100 mg/kg groups. (C-D) Representative images and quantitative analyses of IHC staining for p-STAT3 expression in the (C-a) Con, (C-b) DM, (C-c) DM+PF 25 mg/kg, (C-d) DM+PF 50 mg/kg and (C-e) DM+PF 100 mg/kg groups. (E) Representative Western blots for p-JAK2, JAK2, p-STAT3 and STAT3. (F) Normalized protein expression of p-JAK2/JAK2 and p-STAT3/STAT3 by Western blotting. β -actin was detected as a loading control. All the values are represented as mean \pm SD of at least three independent experiments ($n = 12$ per group). * $p < 0.05$, ** $p < 0.01$ vs. Con; # $p < 0.05$, ## $p < 0.01$ vs. DM. Con, normal control group; DM, diabetic model group; PF, Paconiflorin; IHC, immunohistochemical; NS, not significant. Original magnification, $\times 400$ in A and C.

diabetic kidneys by inhibiting macrophage infiltration and further suppress macrophage-mediated expression of inflammatory factors.

3.3. Paeoniflorin inhibits the activation of JAK2/STAT3 signaling in diabetic kidneys

As described above, Paeoniflorin exhibited an anti-inflammation effect in the diabetic kidney. For this reason, certain molecules and signal pathways associated with inflammation were examined. We initially measured the expression level of p-JAK2 and p-STAT3 proteins in the glomeruli and tubulointerstitium by IHC staining. The present results indicated that the DM group expressed an intense staining of p-JAK2 in the glomeruli and tubulointerstitium area. Furthermore, Paeoniflorin treatment significantly reduced the stained section of p-JAK2 both in glomeruli and tubulointerstitium compared to the DM group in a dose-dependent manner (Figure 4A and 4B). The same results of p-STAT3 expression were detected in the tubulointerstitium area, however, no significant difference was identified for p-STAT3 expression in the glomeruli area among five groups (Figure 4C and 4D).

Consistent with immunohistochemistry analysis, Western blot analysis results further confirmed an upregulation of detectable p-JAK2 and p-STAT3 protein expression levels in the DM group. Compared with the DM group, expression of p-JAK2 and p-STAT3 levels were significantly attenuated by Paeoniflorin treatment for 12 weeks. In contrast, no obvious expression changes of total JAK2 and STAT3 were observed among five groups (Figure 4E and 4F). Thus, these results indicated that the JAK2/STAT3 signaling pathway was activated in DN while Paeoniflorin could exert an inhibition role effect on the JAK2/STAT3 signaling pathway in the STZ-induced diabetic model.

4. Discussion

The global prevalence of DN has been increasing over the last few decades, and satisfactory therapeutic strategies that prevent DN progression still remain a challenge. Renal morphological and functional alterations are two major characterizations of DN. Clinically, microalbuminuria is the earliest detectable biomarker of glomerular damage and is also an independent risk factor for the progression of DN (20). In our study, the KW/BW ratio was used as an indicator of renal hypertrophy index in line with former research (21). Major pathological changes of DN include thickening of glomerular basement membrane, mesangial matrix accumulation and expansion, glomerulosclerosis, tubular interstitial inflammation and fibrosis which eventually lead to proteinuria and renal failure (22).

To date, cumulative studies highlighted beneficial

effects of traditional Chinese medicine (TCM) for delaying DN progression and provided new insights into promising drugs for DN prevention (23). For instance, Paeoniflorin exhibits a significant anti-inflammatory effect in models of rheumatoid arthritis, hepatic fibrosis, and ulcerative colitis (24-26). Notably, research also demonstrated the renoprotective role of Paeoniflorin in kidney disease including renal fibrosis (27) and acute renal injury (28). Thus in this present study, renal protection from Paeoniflorin was evaluated as a novel therapeutic agent for DN. Our findings showed that Paeoniflorin significantly reduced the albuminuria and attenuated renal hypertrophy in STZ-induced diabetic mice. However, Paeoniflorin did not impact the blood glucose levels of diabetic mice which were not consistent with a Fu *et al.* previous study (14). Furthermore, the histology damage in diabetic renal samples was improved a lot after Paeoniflorin treatment. Therefore, our study demonstrated that Paeoniflorin treatment alleviated both the functional and morphological damage of the diabetic mouse model.

Accumulated evidence implied the crucial role of renal inflammation in the development of DN. Recent studies of both animal and human models have showed that macrophage infiltration and the over expression of proinflammatory cytokines lead to renal damage and fibrosis in diabetic kidneys (29). The macrophages associated with renal damage can be divided into: classically activated M1 macrophages involved in inflammation and tissue damage or alternatively activated M2 macrophages mediating tissue repair and renal protection (30). Currently, iNOS is the most widely used specific marker of M1 macrophages (31). In addition, macrophages mediate the over expression levels of proinflammatory cytokines like MCP-1, TNF- α and IL-1 β and finally induce renal damage and fibrosis in human and animal DN models (32). Consistent with our previous studies (15), the renal tissues of the diabetic mice group in our study showed massive CD68-positive macrophage infiltration in the glomerular and tubulointerstitial regions as well as an elevated mRNA expression level of iNOS. Our further investigation showed that expression of pro-inflammatory cytokines including MCP-1, TNF- α and IL-1 β also increased dramatically. We noted that Paeoniflorin treatment apparently reduced macrophage infiltration and iNOS expression levels, thus alleviating the pro-inflammatory cytokines in diabetic mice kidneys. Consequently, our findings suggested the potential anti-inflammatory effect of Paeoniflorin in DN and this protective role probably is associated with the suppression of M1 macrophages infiltration in renal tissues.

Activation of the JAK2/STAT3 signaling pathway has been proven to play a pivotal role in the progression of DN in rodent models and human patients (6). On the basis of recent studies, inhibition of JAK-STAT signaling, including increased expression

of the suppressors of cytokine signaling proteins and pharmacologic inhibition of JAK and STAT proteins, establish a new therapeutic target for DN (33). Baricitinib (ClinicalTrials.gov Identifier NCT01683409), a selective JAK1 and JAK2 inhibitor, has been investigated in phase 2 randomized clinical trials in participants with type 2 diabetic nephropathy (34). The study has proven that baricitinib treatment resulted in a reduction in albuminuria after 3 months' treatment (35). However, the impact of the JAK2/STAT3 signaling pathway in the renoprotective effect of Paeoniflorin has not been examined. Our experiment showed that the significant upregulation of the protein expression of p-JAK2 and p-STAT3 in diabetic kidney decreased a lot after Paeoniflorin intervention. As the activation of the JAK2/STAT3 pathway can be inhibited by Paeoniflorin treatment, the current study demonstrated that the protection of Paeoniflorin against STZ-caused diabetic damage might be *via* suppressing of JAK2/STAT signaling. Therefore, the inhibition of JAK2/STAT3 signals through Paeoniflorin may be a novel method for DN treatment.

As an important signal in inflammation, the activation of the JAK/ STAT pathway had been shown to promote the polarization of M1 macrophages (36). Moreover, Zhai *et al.* (13) revealed that Paeoniflorin reduced M1 cells activity through inhibiting the activation of NF- κ B signaling pathway while enhancing M2 cells function *via* activating STAT6 phosphorylation in mouse bone marrow derived macrophages. Our study along with previous published results proved the anti-inflammatory role of Paeoniflorin in DN. For this reason, we propose a potential theory that Paeoniflorin attenuates renal inflammation in STZ-induced diabetic mice kidney and these protective effects may be associated with the prevention of macrophage activation through inhibiting the JAK2/STAT3 signaling pathway to some extent. However, it remains unclear whether Paeoniflorin prevents macrophage activation *via* the inhibition of JAK2/STAT3 in DN. Therefore, we need further studies to uncover the exact mechanism of Paeoniflorin on JAK2/STAT3 and M1 macrophage activation in DN.

In conclusion, our study confirms that Paeoniflorin treatment alleviates renal lesions by inhibiting the activation of JAK2/STAT3 signaling pathways and the infiltration of macrophages together with the suppression of pro-inflammatory cytokine expression, thus improving the prognosis of DN. Therefore, our findings suggest that Paeoniflorin as a novel and feasible therapeutic strategy to halt the onset and progression of renal inflammation in DN.

Acknowledgements

This current study was supported by the Natural Science Foundation of China (NO. 81374034, 81470965).

References

1. Tuttle KR, Bakris GL, Bilous RW, Chiang JL, de Boer IH, Goldstein-Fuchs J, Hirsch IB, Kalantar-Zadeh K, Narva AS, Navaneethan SD, Neumiller JJ, Patel UD, Ratner RE, Whaley-Connell AT, Molitch ME. Diabetic kidney disease: A report from an ADA Consensus Conference. *Am J Kidney Dis.* 2014; 64:510-533.
2. Zhang L, Long J, Jiang W, Shi Y, He X, Zhou Z, Li Y, Yeung RO, Wang J, Matsushita K, Coresh J, Zhao MH, Wang H. Trends in chronic kidney disease in China. *N Engl J Med.* 2016; 375:905-906.
3. Kanwar YS, Sun L, Xie P, Liu FY, Chen S. A glimpse of various pathogenetic mechanisms of diabetic nephropathy. *Annu Rev Pathol.* 2011; 6:395-423.
4. Duran-Salgado MB, Rubio-Guerra AF. Diabetic nephropathy and inflammation. *World J Diabetes.* 2014; 5:393-398.
5. Fernandez-Fernandez B, Ortiz A, Gomez-Guerrero C, Egido J. Therapeutic approaches to diabetic nephropathy--beyond the RAS. *Nat Rev Nephrol.* 2014; 10:325-346.
6. Brosius FC 3rd, He JC. JAK inhibition and progressive kidney disease. *Curr Opin Nephrol Hypertens.* 2015; 24:88-95.
7. Stark GR, Darnell JE Jr. The JAK-STAT pathway at twenty. *Immunity.* 2012; 36:503-514.
8. Berthier CC, Zhang H, Schin M, Henger A, Nelson RG, Yee B, Boucherot A, Neusser MA, Cohen CD, Carter-Su C, Argetsinger LS, Rastaldi MP, Brosius FC, Kretzler M. Enhanced expression of Janus kinase-signal transducer and activator of transcription pathway members in human diabetic nephropathy. *Diabetes.* 2009; 58:469-477.
9. He DY, Dai SM. Anti-inflammatory and immunomodulatory effects of paeonia lactiflora pall., a traditional chinese herbal medicine. *Front Pharmacol.* 2011; 2:10.
10. Liu HQ, Zhang WY, Luo XT, Ye Y, Zhu XZ. Paeoniflorin attenuates neuroinflammation and dopaminergic neurodegeneration in the MPTP model of Parkinson's disease by activation of adenosine A1 receptor. *Br J Pharmacol.* 2006; 148:314-325.
11. Ouyang J, Xu H, Li M, Dai X, Fu F, Zhang X, Lan Q. Paeoniflorin exerts antitumor effects by inactivating S phase kinase-associated protein 2 in glioma cells. *Oncol Rep.* 2018; 39:1052-1062.
12. Wang C, Yuan J, Wu HX, Chang Y, Wang QT, Wu YJ, Liu LH, Wei W. Paeoniflorin inhibits inflammatory responses in mice with allergic contact dermatitis by regulating the balance between inflammatory and anti-inflammatory cytokines. *Inflamm Res.* 2013; 62:1035-1044.
13. Zhai T, Sun Y, Li H, Zhang J, Huo R, Li H, Shen B, Li N. Unique immunomodulatory effect of paeoniflorin on type I and II macrophages activities. *J Pharmacol Sci.* 2016; 130:143-150.
14. Fu J, Li Y, Wang L, Gao B, Zhang N, Ji Q. Paeoniflorin prevents diabetic nephropathy in rats. *Comp Med.* 2009; 59:557-566.
15. Zhang T, Zhu Q, Shao Y, Wang K, Wu Y. Paeoniflorin prevents TLR2/4-mediated inflammation in type 2 diabetic nephropathy. *Biosci Trends.* 2017; 11:308-318.
16. Breyer MD, Bottinger E, Brosius FC, 3rd, Coffman TM, Harris RC, Heilig CW, Sharma K. Mouse models of diabetic nephropathy. *J Am Soc Nephrol.* 2005; 16:27-45.

17. Zhang LL, Wei W, Wang NP, Wang QT, Chen JY, Chen Y, Wu H, Hu XY. Paeoniflorin suppresses inflammatory mediator production and regulates G protein-coupled signaling in fibroblast-like synoviocytes of collagen induced arthritic rats. *Inflamm Res.* 2008; 57:388-395.
18. Zhang H, Li P, Burczynski FJ, Gong Y, Choy P, Sha H, Li J. Attenuation of Diabetic Nephropathy in Otsuka Long-Evans Tokushima Fatty (OLETF) Rats with a Combination of Chinese Herbs (Tangshen Formula). *Evid Based Complement Alternat Med.* 2011; 2011:613737.
19. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2⁻(-Delta Delta C(T)) Method. *Methods.* 2001; 25:402-408.
20. Bakris GL, Molitch M. Microalbuminuria as a risk predictor in diabetes: The continuing saga. *Diabetes care.* 2014; 37:867-875.
21. Shuhong Z, Hongjun L, Bo C, Li X, Bingyin S. The effect of dantonic pill on beta -catenin expression in a rat model of streptozotocin-induced early stage of diabetic nephropathy. *J Diabetes Res.* 2013; 2013:848679.
22. Dalla Vestra M, Saller A, Mauer M, Fioretto P. Role of mesangial expansion in the pathogenesis of diabetic nephropathy. *J Nephrol.* 2001; 14 Suppl 4:S51-S57.
23. Xue R, Gui D, Zheng L, Zhai R, Wang F. Mechanistic insight and management of diabetic nephropathy: Recent progress and future perspective. *J Diabetes Res.* 2017; 2017:1839809.
24. Gu P, Zhu L, Liu Y, Zhang L, Liu J, Shen H. Protective effects of paeoniflorin on TNBS-induced ulcerative colitis through inhibiting NF-kappaB pathway and apoptosis in mice. *Int Immunopharmacol.* 2017; 50:152-160.
25. Jia Z, He J. Paeoniflorin ameliorates rheumatoid arthritis in rat models through oxidative stress, inflammation and cyclooxygenase 2. *Exp Ther Med.* 2016; 11:655-659.
26. Zhao Y, Ma X, Wang J, Zhu Y, Li R, Wang J, He X, Shan L, Wang R, Wang L, Li Y, Xiao X. Paeoniflorin alleviates liver fibrosis by inhibiting HIF-1alpha through mTOR-dependent pathway. *Fitoterapia.* 2014; 99:318-327.
27. Zeng J, Dou Y, Guo J, Wu X, Dai Y. Paeoniflorin of *Paeonia lactiflora* prevents renal interstitial fibrosis induced by unilateral ureteral obstruction in mice. *Phytomedicine.* 2013; 20:753-759.
28. Wang P, Wang W, Shi Q, Zhao L, Mei F, Li C, Zuo T, He X. Paeoniflorin ameliorates acute necrotizing pancreatitis and pancreatitis induced acute renal injury. *Mol Med Rep.* 2016; 14:1123-1131.
29. Ma J, Wu H, Zhao CY, Panchapakesan U, Pollock C, Chadban SJ. Requirement for TLR2 in the development of albuminuria, inflammation and fibrosis in experimental diabetic nephropathy. *Int J Clin Exp Pathol.* 2014; 7:481-495.
30. Meng XM, Tang PM, Li J, Lan HY. Macrophage Phenotype in Kidney Injury and Repair. *Kidney Dis (Basel).* 2015; 1:138-146.
31. Guiteras R, Flaquer M, Cruzado JM. Macrophage in chronic kidney disease. *Clin Kidney J.* 2016; 9:765-771.
32. Wu CC, Sytwu HK, Lin YF. Cytokines in diabetic nephropathy. *Adv Clin Chem.* 2012; 56:55-74.
33. Brosius FC, Tuttle KR, Kretzler M. JAK inhibition in the treatment of diabetic kidney disease. *Diabetologia.* 2016; 59:1624-1627.
34. Lacava V, Pellicano V, Ferrajolo C, Cernaro V, Visconti L, Conti G, Buemi M, Santoro D. Novel avenues for treating diabetic nephropathy: New investigational drugs. *Expert Opin Investig Drugs.* 2017; 26:445-462.
35. Heerspink HJ, De Zeeuw D. Novel anti-inflammatory drugs for the treatment of diabetic kidney disease. *Diabetologia.* 2016; 59:1621-1623.
36. Zhou D, Huang C, Lin Z, Zhan S, Kong L, Fang C, Li J. Macrophage polarization and function with emphasis on the evolving roles of coordinated regulation of cellular signaling pathways. *Cell Signal.* 2014; 26:192-197.

(Received January 27, 2018; Revised March 31, 2018; Accepted April 11, 2018)

Effects of three forms of local anesthesia on perioperative fentanyl-induced hyperalgesia

Lu Chang¹, Fang Ye², Quehua Luo¹, Zewen Wang¹, Yimin Wang¹, Zhengyuan Xia^{1,3}, Haihua Shu^{1,*}

¹Department of Anesthesiology, Guangdong Second Provincial General Hospital, Guangzhou, Guangdong, China;

²Department of Anesthesiology, First Affiliated Hospital of Sun Yat-sen University, Guangzhou, Guangdong, China;

³Department of Anesthesiology, The University of Hong Kong, Hong Kong SAR, China.

Summary

Both local infiltration analgesia (LIA) and nerve block are common analgesic modalities for pain relief after surgery. The aim of the current study was to investigate the effects of those two modalities on pain behavior and the expression of pro-inflammatory cytokines such as interleukin (IL)-1 β and IL-6 and tumor necrosis factor- α (TNF- α) in the spinal cord and dorsal root ganglion (DRG) in a rat model of perioperative fentanyl induced hyperalgesia. Rats were injected with fentanyl (60 μ g/kg) 4 times and received a plantar incision after the second injection or they received pre-incision LIA and sciatic nerve block (SNB) or post-incision LIA with levobupivacaine (0.5%, 0.2 mL). Mechanical and thermal nociceptive thresholds were assessed using the tail pressure test and paw withdrawal test on the day before drug injection, 1 and 4 hours after injection, and 1-7 days later. The lumbar spinal cord and dorsal root ganglia were collected from rats in each group to measure IL-1 β , IL-6, and TNF- α on the day before drug injection, 4 hours after injection, and 1, 3, 5, and 7 days later. Fentanyl and an incision induced a significantly delayed mechanical hyperalgesia in the tail and thermal hyperalgesia in both hind paws and up-regulation of pro-inflammatory cytokines in the spinal cord and dorsal root ganglia. Rats treated with pre-incision LIA and SNB or post-incision LIA had alleviated hyperalgesia and significantly reduced levels of IL-1 β , IL-6, and TNF- α compared to the control group. LIA and SNB partly prevented perioperative fentanyl-induced hyperalgesia and up-regulation of pro-inflammatory cytokines in the spinal cord and dorsal root ganglia.

Keywords: Regional anesthesia, hyperalgesia, spinal cord, dorsal root ganglion, pro-inflammatory cytokines

1. Introduction

Postsurgical pain is a common complication of many surgical procedures and the most frequent cause of discomfort after surgery (1). Cutting and handling tissue during surgery causes trauma and inflammation that in turn activate nociceptors and induce postsurgical pain (2).

Opioids, used perioperatively as a basal analgesic, cause opioid-induced hyperalgesia (OIH) that may aggravate postsurgical pain (3).

Peripheral nerve block, such as local infiltration anesthesia (LIA) and regional nerve block, can enhance perioperative pain control and may have more positive effects (4,5). To the extent known, however, previous studies have rarely simultaneously investigated the effect of local anesthesia on a surgery-induced increase in pro-inflammatory cytokines in the spinal cord and dorsal root ganglia. Since the activation of spinal glia and upregulation of pro-inflammatory cytokines are critical to the development of post-surgical pain (6,7), the aim of the current study was to investigate the effects of LIA and sciatic nerve block (SNB) on pain behavior and

Released online in J-STAGE as advance publication April 15, 2018.

*Address correspondence to:

Dr. Haihua Shu, Department of Anesthesiology, Guangdong Second Provincial General Hospital, 466# Xingang Middle Road, Guangzhou, Guangdong 510317, China.

E-mail: shuhaihua@hotmail.com

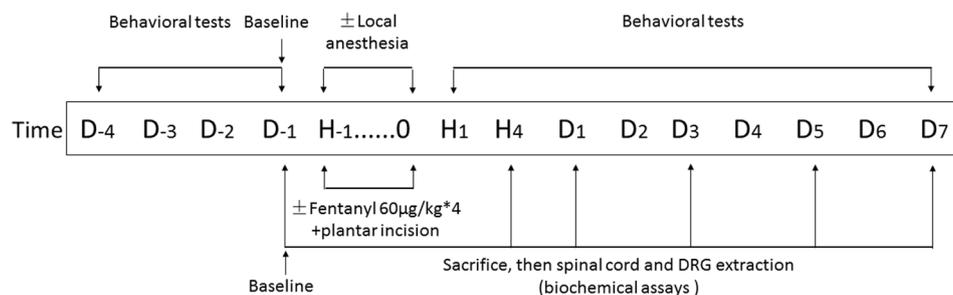


Figure 1. Medication and Measurement Protocol. Rats were randomly assigned into 5 groups ($n = 28$ in each group). Rats in the normal group received no treatment. The other 4 groups of rats were subcutaneously (s.c.) injected with fentanyl at a dose of $60 \mu\text{g}/\text{kg}$ 4 times at 15-minute intervals and received a plantar incision after the second injection. Rats in 3 groups (treatment groups) received pre-incision local infiltration anesthesia (LIA) or sciatic nerve block (SNB) or post-incision LIA with 0.2 ml of 0.5% levobupivacaine respectively. Behavioral tests were performed daily for 11 days (D-4 to D-1 and D1 to D7) and 1 and 4 hours (H1 and H4) after treatment. The thermal and mechanical thresholds on D-1 were recorded to serve as baseline values. The lumbar spinal cord (L3-L6) and both ipsilateral and contralateral dorsal root ganglia (L3-L6) were collected on D-1, H4, D1, D3, D5, and D7 to examine the expression of interleukin- 1β (IL- 1β), interleukin-6 (IL-6), and tumor necrosis factor (TNF- α).

the expression of pro-inflammatory cytokines such as interleukin (IL)- 1β and IL-6 and tumor necrosis factor- α (TNF- α) in the spinal cord and dorsal root ganglia of rats. Given the fact that opioids are indispensable analgesics in most surgeries, this study used a rat model of perioperative fentanyl-induced hyperalgesia.

2. Materials and Methods

2.1. Animals

This study was approved by the Institutional Laboratory Animal Care and Use Committee of Sun Yat-sen University and was carried out in accordance with the guidelines of the Chinese Physiological Society regarding the care of experimental animals. Adult male Sprague Dawley rats weighing $200\text{-}250 \text{ g}$ (purchased from Sun Yat-sen University, Guangzhou China) at the start of the experiments were used. All rats were housed 4 animals per cage in a specific pathogen-free environment at $25 \pm 2 \text{ }^\circ\text{C}$ with a 12-hour light/dark cycle. Food and water were available ad libitum. Rats were randomly assigned into 5 groups ($n = 28$ per group), including a normal group (N), a fentanyl and a plantar incision (FI) group, a pre-incision LIA (Pre-L) group, a pre-incision SNB (Pre-S) group, and a post-incision LIA (Post-L) group.

2.2. Medication and measurement protocol

Rats in the normal group received no intervention. The four other groups of rats received subcutaneous injections of fentanyl ($60 \mu\text{g}/\text{kg}$) 4 times at 15-minute intervals and underwent surgery after the second injection. Rats in the Pre-L group and the Pre-S group received LIA or an SNB with a single dose of levobupivacaine (0.2 mL , 0.5%) before a plantar incision and rats in the Post-L group received LIA after a plantar incision. Behavioral tests as described below were performed daily for 11 days (D-4 to D-1, D1 to D7) and 1 and 4 hours (H1 and H4) after

fentanyl injections on the day of treatment. Behavioral tests were performed on D-4 to D-2 to acclimate rats to the apparatus and environment. Values on D-1 before treatment were recorded as baseline values. All behavioral tests were performed by a single investigator blinded to the treatment groups. Each test (between 9:00 and 14:00) was repeated 3 times at 3-minute intervals, and results were averaged to obtain a final value. The lumbar spinal cord (L3-L6) and dorsal root ganglia (L3-L6) were collected on D-1, H4, D1, D3, D5, and D7 to examine the expression of IL- 1β , IL-6, and TNF- α (Figure 1).

2.3. Surgery and local SNB

2.3.1. Surgery

Rats in the FI, Pre-L, Post-L, and Pre-S groups were anesthetized with 1.5% to 2% isoflurane (Hebei Jiupai Pharmaceutical Co., Hebei, China). A 1-cm incision was made through the skin, fascia, and muscle of the plantar aspect of the left hind paw after skin disinfection. The incision was started 0.5 cm from the proximal edge of the heel and extended to the toes (8). A 4-0 suture was used to suture the skin after the wound was dabbed with a cotton swab to stop bleeding. All incisions were made after the second injection of fentanyl.

2.3.2. LIA

Rats in the Pre-LIA group and the Post-LIA group respectively received LIA on the left before or after a plantar incision following anesthesia with 1.5% to 2% isoflurane. The needle was inserted about 4 mm , and levobupivacaine (0.2 mL , 0.5%) was administered slowly over 5 seconds when no bleeding occurred.

2.3.3. SNB

Rats in the Pre-S group received an SNB on the left

before a plantar incision after anesthesia with 1.5% to 2% isoflurane. The greater trochanter was located by palpation in the left hind limb, and a needle (26G) was introduced slowly from the left femoral shaft into the sciatic notch. If slight contraction, shaking, or spasming of the left leg was noted, the needle was deemed to be close to the sciatic nerve (5). To avoid nerve damage, the needle was pulled out about 1 mm, and levobupivacaine (0.2 mL, 0.5%) was administered slowly over 5 seconds when no bleeding occurred. Rats with the toes and foot plantar flexed with no splaying 5-10 minutes after the SNB and without dyskinesia in the left hind limb 1 day after SNB were used in experiments.

2.4. Behavior tests

2.4.1. Tail pressure test

The mechanical threshold of a response in the tail was evaluated with an electronic device (ZH-YLS-3E, Anhui Zhenghua Biological Equipment Co., China) as described previously (9). Each rat was immobilized in a cylindrical immobilization device, and a mark was made 10 cm from the tip of the tail. The marked portion of the tail was placed on a platform where a plastic wedge applied pressure to the surface of the tail. The test was stopped and the value was recorded as the mechanical nociceptive threshold (MNT) if the rat flicked its tail, screamed, or struggled. A cutoff pressure of 600 g was used to avoid tissue damage.

2.4.2. Paw withdrawal test

Thermal nociception was measured with a plantar test apparatus (IITC, USA) as described previously (10). Each rat was put in a plexiglass box on a glass plate. A beam of light (30 W, 5 mm) passed through the glass to the sole near the toes both ipsilateral and contralateral to the surgical site. The paw withdrawal latency (PWL) was recorded as the time from the projection of the light

beam to paw withdrawal. A cut-off time of 20 seconds was used to avoid tissue damage.

2.5. Enzyme-linked Immunosorbent Assay (ELISA)

The lumbar spinal cord (L3-L6) and dorsal root ganglia (L3-L6) were collected from each rat in a 2-mL EP tube filled with iced PBS (weight/volume ratio: 1 mg/100 μ L). After the sample was homogenized and centrifuged, the supernatants were removed for immediate use. The amounts of IL-1 β , TNF- α , and IL-6 were assayed using commercially available rat-specific ELISA kits (Shanghai Yikesai Biological Products Co., Shanghai, China) in accordance with the manufacturer's instructions as described previously (Ref. PMID: 27195494).

2.6. Statistical analysis

All data are presented as the mean \pm standard deviation (SD). SPSS for Windows 20.0 (SPSS Inc, Chicago, Illinois, USA) was used for statistical analysis. Changes in continuous variables over time were analyzed with a repeated measures analysis of variance, followed by Fisher's protected least significant difference (PLSD) test. Differences among groups at each time point and differences in each group at multiple time points were analyzed using an analysis of variance and the PLSD test. A paired samples *t* test was used for comparison (incised ipsilateral site vs. contralateral site). *p* < 0.05 was considered statistically significant. A scatter diagram was drawn before statistical analysis to exclude apparent outliers.

3. Results

3.1. LIA and SNB partly prevented mechanical and thermal hyperalgesia induced by fentanyl and a plantar incision

Changes in MNTs are shown in Figure 2A. The MNTs

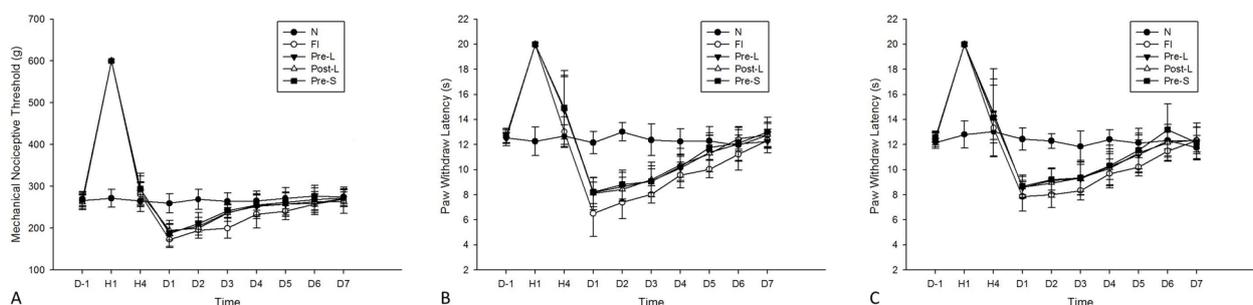


Figure 2. Local infiltration analgesia (LIA) and sciatic nerve block (SNB) partly prevent mechanical and thermal hyperalgesia induced by fentanyl and a plantar incision. SD rats in the normal group received no intervention. The other rats received subcutaneous fentanyl (60 μ g/kg*4) and a plantar incision with or without local anesthesia with levobupivacaine (0.5%, 0.2 mL). (N group = normal rats; FI group = rats receiving fentanyl and an incision; Pre-L group = rats receiving pre-incision LIA, Post-L group = rats receiving post-incision LIA, and Pre-S group = rats receiving pre-incision SNB). The mechanical nociceptive threshold (g) (A) and paw withdrawal latency (s) of incised paws (B) and contralateral paws (C) on the day before drug injection (D-1), 1 and 4 hours after injection, and 1, 2, 3, 4, 5, 6, and 7 days later (D1-D7) are presented as the mean \pm SD (*n* = 8).

increased significantly after 1 hour in the FI and Post-L groups and after 1 and 4 hours in the Pre-L and Pre-S groups, and the MNTs decreased significantly on days 1-5 after fentanyl injections. The decrease in MNTs in the FI group continued for 5 days (D1 to D5) and the decrease in MNTs in the Pre-L, Post-L, and Pre-S groups continued for 3 days (D1 to D3) compared to baseline values and MNTs in the normal group. In addition, the MNTs in the FI group were significantly lower than those in the Pre-L group on D3, in the Post-L group on days 1 and 3, and in the Pre-S group on days 3 and 4 ($p < 0.05$ according to a PLSD test).

The changes in thermal nociceptive latencies (shown as PWLs) ipsilateral to the surgical site are shown in Figure 2B. Similar to MNTs, the PWLs increased significantly after 1 hour in the FI group and after 1 and 4 hours in the Pre-L, Post-L, and Pre-S groups, and the PWLs decreased significantly on days 1-6 after fentanyl injections. The decrease in PWLs continued for 6 days (D1 to D6) in the FI group, for 4 days (D1 to D4) in the Pre-L and Pre-S groups, and for 5 days (D1 to D5) in the Post-L group compared to baseline values and PWLs in the normal group. In addition, the PWLs in the FI group were significantly lower than those in Pre L on days 1, 2 and 5, in the Post-L group on days 1, 3, 5 and 6, in the Pre-S group on days 1, 2 and 5 ($p < 0.05$ according to a PLSD test).

Interestingly, the changes in PWLs at sites contralateral to the surgical site (Figure 2C) were similar to those at ipsilateral sites (Figure 2B). Injections of fentanyl produced significant early thermal analgesia and subsequently produced significant hyperalgesia on days 1-5 in the FI, Pre-L, Post-L, and Pre-S groups. However, the thermal hyperalgesia exhibited at contralateral sites lasted less time than that at incision sites in the FI group (5 days vs. 6 days) and in the Post-L group (4 days vs. 5 days). Moreover, the PWLs at ipsilateral sites were significantly lower than those at contralateral sites on days 1, 2 and 3 in the FI group and on days 1 and 2 in the Pre-L, Post-L, and Pre-S groups after surgery ($p < 0.05$ for all according to a paired-samples t test).

No statistically significant differences in MNTs and PWLs were noted among the Pre-L, Post-L, and Pre-S groups.

In brief, the above results suggested that LIA and SNB partly prevented mechanical and thermal hyperalgesia induced by fentanyl and an incision.

3.2. LIA and SNB partly prevented an increase in pro-inflammatory cytokines in the spinal cord and dorsal root ganglia induced by fentanyl and a plantar incision

Levels of IL-1 β (Figure 3), IL-6 (Figure 4), and TNF- α (Figure 5) expression in the spinal cord and dorsal root ganglia were determined 1 day before fentanyl injections, 4 hours after injections, and 1, 3, 5, and 7 days afterwards.

In the normal group, the expression of IL-1 β , IL-6 and TNF- α did not significantly change over time. However, the level of IL-1 β significantly increased in the spinal cord and dorsal root ganglia in the FI, Pre-L, Post-L, and Pre-S groups 1, 3, 5, and 7 days after fentanyl injections compared to baseline values and the levels of IL-1 β in the normal group ($p < 0.05$ for all according to a PLSD test). Moreover, the levels of IL-1 β in the FI group were significantly higher than those in the Pre-L group on days 1 and 3, in the Post-L group on days 1, 3, and 5, and in the Pre-S group on days 1, 3, 5, and 7 after fentanyl injections in the spinal cord ($p < 0.05$ according to a PLSD test) (Figure 3A). The levels of IL-1 β in the FI group were significantly higher than those in the Pre-L and Pre-S groups on days 5 and 7, and in the Post-L group on day 5 after fentanyl injections in the ipsilateral dorsal root ganglion (DRG) ($p < 0.05$ for all according to a PLSD test) (Figure 3B). Similarly, the levels of IL-1 β in the FI group were significantly higher than those in the Pre-L and Pre-S groups on day 5 and in the Post-L group on days 1 and 7 after fentanyl injections in the contralateral DRG ($p < 0.05$ according to a PLSD test) (Figure 3C). However, the levels of IL-1 β in the ipsilateral DRG were significantly higher than those in the contralateral DRG in the FI and Pre-L groups 4 hours after injection and 1, 3, 5, and 7 days later, in the Post-L group 4 hours after injection and 1 and 3 days later, and in the Pre-S group 4 hours after injection and 1, 3, and 5 days later ($p < 0.05$ according to a paired samples t test).

Similarly, the levels of IL-6 significantly increased in the FI, Pre-L, Post-L, and Pre-S groups on days 1, 3, 5, and 7 in the spinal cord, in the FI and Pre-S groups on days 1, 3, 5, and 7, and in the Pre-L and Post-L groups on days 3, 5, and 7 in the dorsal root ganglia after fentanyl injections compared to baseline values and the levels of IL-6 in the normal group ($p < 0.05$ according to a PLSD test). Moreover, the levels of IL-6 in the FI group were significantly higher than those in the Pre-L and Post-L groups on days 1, 3 and 5 and in the Pre-S group on days 1, 3, 5, and 7 after fentanyl injections in the spinal cord ($p < 0.05$ according to a PLSD test) (Figure 4A). The levels of IL-6 in the FI group were significantly higher than those in the Pre-L and Post-L groups on days 5 and 7 and in the Pre-S group on days 3, 5, and 7 after fentanyl injections in the ipsilateral DRG ($p < 0.05$ according to a PLSD test) (Figure 4B). The levels of IL-6 in the FI group were significantly higher than those in the Pre-L group on days 3, 5, and 7, in the Post-L group on days 5 and 7, and in the Pre-S group on day 5 after fentanyl injections in the contralateral DRG ($p < 0.05$ according to a PLSD test) (Figure 4C). However, the levels of IL-6 in the ipsilateral DRG were significantly lower than those in the contralateral DRG on days 5 and 7 in the Pre-L group and on day 5 in the Post-L group ($p < 0.05$ according to a paired samples t test).

Similarly, the levels of TNF- α significantly

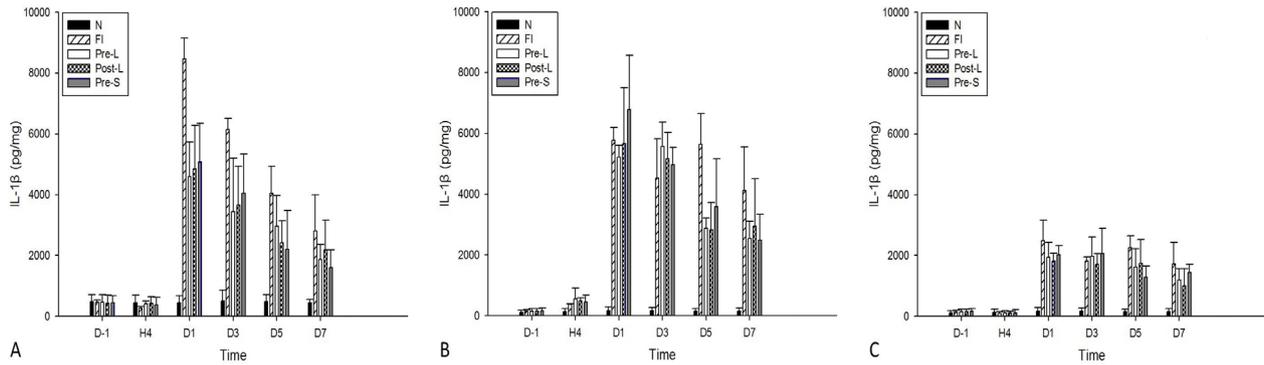


Figure 3. Local infiltration analgesia (LIA) and sciatic nerve block (SNB) partly prevent an increase in interleukin 1 β (IL-1 β) in the spinal cord induced by fentanyl and a plantar incision. SD rats in the normal group received no intervention. The other rats received subcutaneous fentanyl (60 μ g/kg*4) and a plantar incision with or without local anesthesia with levobupivacaine (0.5%, 0.2 mL). (N group = normal rats; FI group = rats receiving fentanyl and an incision; Pre-L group = rats receiving pre-incision LIA, Post-L group = rats receiving post-incision LIA, and Pre-S group = rats receiving pre-incision SNB). The lumbar spinal cord (A) and the ipsilateral DRG (B) and contralateral DRG (C) were collected from surgical sites in 4 rats from each group. The expression of IL-1 β (pg/mg) on the day before drug injection (D-1), 4 hours after injection (H4), and 1, 3, 5, and 7 days later (D1, D3, D5, and D7) was detected with an enzyme-linked immunosorbent assay (ELISA). All data are presented as the mean \pm SD ($n = 4$).

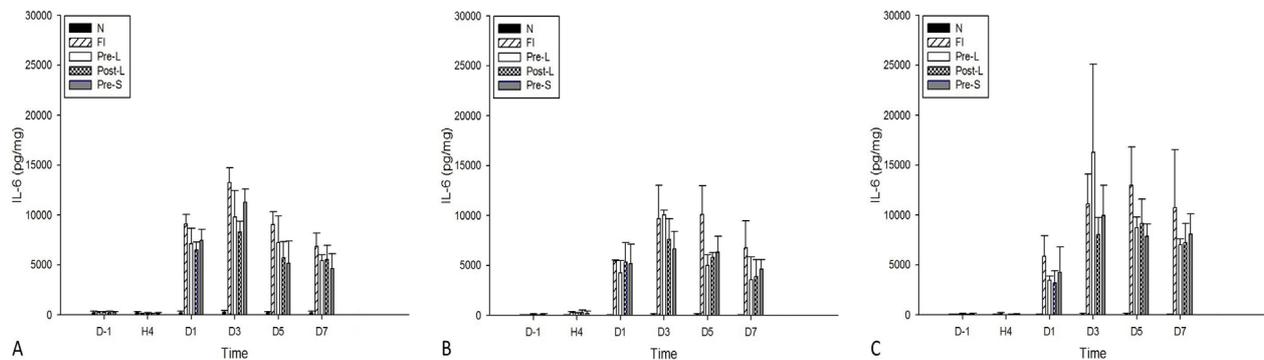


Figure 4. Local infiltration analgesia (LIA) and sciatic nerve block (SNB) partly prevent an increase in interleukin 6 (IL-6) in the spinal cord and dorsal root ganglia induced by fentanyl and a plantar incision. SD rats in the normal group received no intervention. The other rats received subcutaneous fentanyl (60 μ g/kg*4) and a plantar incision with or without local anesthesia with levobupivacaine (0.5%, 0.2 mL). (N group = normal rats; FI group = rats receiving fentanyl and an incision; Pre-L group = rats receiving pre-incision LIA, Post-L group = rats receiving post-incision LIA, and Pre-S group = rats receiving pre-incision SNB). The lumbar spinal cord (A) and the ipsilateral DRG (B) and contralateral DRG (C) were collected from surgical sites in 4 rats from each group. The expression of IL-6 (pg/mg) on the day before drug injection (D-1), 4 hours after injection (H4), and 1, 3, 5, and 7 days later (D1, D3, D5, and D7) was detected with an enzyme-linked immunosorbent assay (ELISA). All data are presented as the mean \pm SD ($n = 4$).

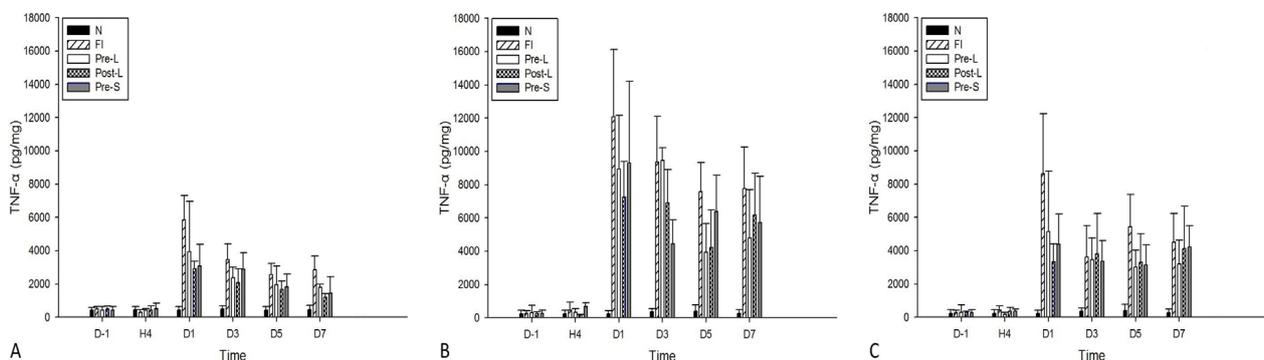


Figure 5. Local infiltration analgesia (LIA) and sciatic nerve block (SNB) partly prevent an increase in tumor necrosis factor α (TNF- α) in the spinal cord and dorsal root ganglia induced by fentanyl and a plantar incision. SD rats in the normal group received no intervention. The other rats received subcutaneous fentanyl (60 μ g/kg*4) and plantar incision with or without local anesthesia with levobupivacaine (0.5%, 0.2 mL). (the N group = normal rats; FI group = rats receiving fentanyl and an incision; the Pre-L group = rats receiving pre-incision LIA, the Post-L group = rats receiving post-incision LIA, the Pre-S group = rats receiving pre-incision SNB). The lumbar spinal cord (A) and the ipsilateral DRG (B) and contralateral DRG (C) were collected from surgical sites in 4 rats from each group. The expression of TNF- α (pg/mg) on the day before drug injection (D-1), 4 hours after injection (H4), and 1, 3, 5, and 7 days later (D1, D3, D5, and D7) was detected with an enzyme-linked immunosorbent assay (ELISA). All data are presented as the mean \pm SD ($n = 4$).

increased in the FI and Pre-L groups on days 1, 3, 5, and 7 in the spinal cord and dorsal root ganglia, in the Post-L and Pre-S groups on days 1, 3 and 5 in the spinal cord, and in the Post-L and Pre-S groups on days 1, 3, 5, and 7 in the dorsal root ganglia after fentanyl injections compared to baseline values and the levels of TNF- α in the normal group ($p < 0.05$ for all according to a PLSD test). Moreover, the levels of TNF- α in the FI group were significantly higher than those in the Pre-L group on day 1, in the Post-L group on days 1, 3 and 7, and in the Pre-S group on days 1 and 7 after fentanyl injections in the spinal cord ($p < 0.05$ according to a PLSD test) (Figure 5A). The levels of TNF- α in the FI group were significantly higher than those in the Pre-L group on days 1, 5, and 7, in the Post-L group on days 1 and 5, and in the Pre-S group on day 3 after fentanyl injections in the ipsilateral DRG ($p < 0.05$ for all according to a PLSD test) (Figure 5B). The levels of TNF- α in the FI group were significantly higher than those in the Pre-L, Post-L, and Pre-S groups on days 1 and 5 after fentanyl injections in the contralateral DRG ($p < 0.05$ according to a PLSD test) (Figure 5C). However, the levels of TNF- α in the ipsilateral DRG were significantly higher than those in the contralateral DRG on day 3 in the FI and Pre-L groups and on day 1 in the Post-L group ($p < 0.05$ according to a paired samples t test).

In brief, the above results indicated that local anesthesia partly prevented an increase in IL-1 β , IL-6, and TNF- α in the spinal cord and dorsal root ganglia induced by fentanyl and a plantar incision. Fentanyl and an incision induced a greater increase in IL-1 β and TNF- α but a lower increase in IL-6 in the ipsilateral DRG than in the contralateral DRG.

4. Discussion

This study found that an incision and fentanyl induced mechanical and thermal nociceptive hyperalgesia and increases in the pro-inflammatory cytokines IL-1 β , IL-6, and TNF- α in the spinal cord and dorsal root ganglia ipsilateral and contralateral to the surgical site in rats. Local anesthesia with levobupivacaine partly prevented perioperative fentanyl-induced hyperalgesia and up-regulation of pro-inflammatory cytokines in the spinal cord and dorsal root ganglia. Three forms of local anesthesia - pre-incision LIA, pre-incision SNB, and postoperative LIA - had a similar effect on postoperative pain.

A previous study by the current authors (11) and another study (12) found that opioids such as fentanyl induce significant hyperalgesia, especially when a high dose has been administered. This phenomenon is called opioid-induced hyperalgesia (OIH). Although the detailed mechanism of this phenomenon is still unclear, it is presumably associated with activation of the N-methyl-D-aspartate (NMDA) receptor, reduced facilitation mediated by on- and off- cells within

the rostro-ventral medulla (RVM), up-regulation of spinal dynorphin, an increase in excitatory peptide neurotransmitters such as cholecystokinin (CCK), and activation of the transient receptor potential vanilloid receptor (TRPV1) (13). Neuroimmune and neuroinflammatory interactions in the central nervous system (CNS) have also been found to contribute to the development of OIH (11,14). A previous study by the current authors found that repeated injections of high doses of fentanyl (60 $\mu\text{g}/\text{kg}^*4$) induced a significant decrease in the mechanical and thermal threshold, activation of microglia in the spinal cord, and increased pro-inflammatory cytokines such as IL-1 β , TNF- α , and IL-6 in the spinal cord and DRG (11).

A surgical incision and trauma usually induce postoperative mechanical allodynia and thermal hyperalgesia in human and animal models (15,16). Interestingly, a previous study by the current authors (11) found that intraoperative administration of high-dose fentanyl resulted in greater hyperalgesia than a surgical incision only, as other studies had found (17,18). These findings suggest that perioperative opioids aggravate surgery-induced postoperative hyperalgesia, and especially at high doses; this finding is consistent with the results of several clinical trials (3,13).

The current study investigated mechanical nociception in the tail and thermal nociception at sites ipsilateral and contralateral to a surgical incision, unlike most studies in which the authors assessed only hyperalgesia at the surgical site (17,18). Surprisingly, mechanical hyperalgesia in the tail and thermal hyperalgesia at contralateral sites were significant and obvious. To be fair, thermal hyperalgesia at contralateral sites was less severe than that at ipsilateral sites early after surgery (Figures 2B and 2C). Since a tail pressure test reflects spinal reflexes involving both sides of the spinal cord (19) and since significant hyperalgesia was noted on the contralateral side, the current findings suggest that central sensitization may largely be involved in the development of delayed hyperalgesia induced by surgery and intraoperative fentanyl.

The current study also investigated pro-inflammatory cytokines such as IL-1 β , IL-6, and TNF- α in the spinal cord and dorsal root ganglia. Results indicated that they significantly increased in the spinal cord and dorsal root ganglia, following a surgical intervention and intraoperative administration of fentanyl with or without local anesthesia. The conjecture is that the mechanism of central sensitization mainly contributed to the hyperalgesia induced by surgery with fentanyl. However, the up-regulation of pro-inflammatory cytokines in the dorsal root ganglia suggests that a peripheral mechanism has also contributed to hyperalgesia to some extent. As an example, a greater increase in IL-1 β and TNF- α in the DRG ipsilateral to the surgical site than in the contralateral DRG may explain why thermal hyperalgesia at ipsilateral sites was more severe than that

at contralateral sites.

Local anesthesia, including LIA and regional nerve block, has been found to prevent the development of postoperative pain at least during the acute postoperative period (4,20,21). The inhibitory effect on postoperative pain by local anesthesia may be mediated by preventing the transmission of nociceptive stimuli from the peripheral to the central nervous system. Clinically, opioids are commonly used during surgical procedures. Therefore, patients are usually given opioids for local anesthesia. The question is whether local anesthesia can prevent hyperalgesia induced by both a surgical incision and an opioid. The current study found that three forms of local anesthesia (including pre- and post-operative LIA and SNB) partly but clearly relieved mechanical hyperalgesia in the tail and thermal hyperalgesia at sites ipsilateral and contralateral to a surgical incision as a result of a plantar incision and perioperative fentanyl. These results corroborate the contention of Rivat *et al.* (22) that regional anesthesia is able to reduce postoperative acute hyperalgesia and chronic pain induced by the surgery itself and intraoperative OIH (22). However, the finding disagrees with the results of Meleine *et al.*, who found that SNB failed to prevent incisional pain when high-dose fentanyl was administered to rats (23). The reason for this inconsistency is unclear. Meleine *et al.* postulated that high doses of fentanyl for intraoperative analgesia induced central sensitization and that SNB could not prevent central sensitization (23). However, the current study found that local anesthesia partly prevented mechanical and thermal hyperalgesia distant from a surgical site, and this study noted an increase in pro-inflammatory cytokines in the spinal cord and ipsilateral DRG induced by an incision and fentanyl in rats (Figure 2A,C and Figure 3,4,5). This indicates that local anesthesia may at least partly prevent central sensitization. The mechanism by which local anesthesia inhibits hyperalgesia distant from a surgical site is unclear. A previous study found that changes in the excitability of sensory neurons in the DRG may generate central connections to the spinal cord and then activate the glia in the spinal cord (24). The activation of spinal glia and an increase in pro-inflammatory cytokines might induce a secondary increase in pro-inflammatory cytokines in the contralateral DRG. This is perhaps part of the reason why local infiltration anesthesia and a nerve block at the ipsilateral site ultimately had an effect on hyperalgesia distant from a surgical site.

Chronic pain is associated with long-term changes in levels of pro-inflammatory cytokines in the spinal cord and the excitability of sensory neurons in the DRG (25). These findings imply that uncontrolled and persistent activity of pro-inflammatory cytokines such as IL-1 β , IL-6, and TNF- α may induce a transition from acute pain to chronic pain. This suggests that the anti-inflammatory

benefits of local anesthesia were a result of reducing short-term postoperative pain as well as preventing the development of chronic postoperative pain somewhat. In addition, sufficient local anesthesia can also facilitate a lower dosage of opioids perioperatively, which may reduce the incidence of OIH.

In the current study, pre-incision LIA and a pre-incision SNB had similar effectiveness in terms of anti-nociceptive and anti-inflammatory activity. Pre-incision LIA has the advantages of a simpler procedure and it causes fewer adverse events, so it is warranted for pain control. Moreover, pre-incision LIA and post-incision LIA differed little in terms of anti-nociceptive and anti-inflammatory activity, indicating that remedial local anesthesia is still valid.

The current study has several limitations. *i)* The dose of fentanyl given to rats was 240 $\mu\text{g}/\text{kg}$ in total, which is equal to 38 $\mu\text{g}/\text{kg}$ in humans after converting based on body surface area (26). This dose was much higher than the dose commonly used in clinical practice. High-dose fentanyl was used in the rat model because a previous study by the current authors noted behavioral hyperalgesia and spinal inflammation induced by fentanyl. The higher dose of fentanyl may cause more distinct behavior and neuroinflammation and allow those mechanisms to be more clearly understood. However, the clinical significance and relevance of this study diminished since such a high dose of fentanyl may cause numerous adverse effects, which is why it is not usually administered to patients undergoing surgery. Further preclinical studies using animal models need to examine surgery and clinical doses of fentanyl in order to obtain clinically relevant data and to facilitate the transition from bench to bed. *ii)* This study used an animal model combining a surgical incision with intraoperative fentanyl. This approach may be similar to clinical practice and provide useful data for clinical reference since opioids are indispensable during surgery. However, this approach incorporates the effect of local anesthesia on hyperalgesia and neuroinflammation, and it does not allow a distinction between fentanyl-induced hyperalgesia and post-incision hyperalgesia since both conditions can induce hyperalgesia.

In conclusion, the current study noted increased expression of pro-inflammatory cytokines such as IL-1 β , IL-6, and TNF- α in the spinal cord and dorsal root ganglia in a rat model of a plantar incision with intraoperative fentanyl. Pre-incision LIA, a pre-incision SNB, or post-incision LIA can partly prevent hyperalgesia and up-regulation of pro-inflammatory cytokines in the spinal cord and dorsal root ganglia induced by surgery and fentanyl.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (project no. 81571071) and

the Guangdong Provincial Natural Science Foundation of China (project no. 2014A030313203).

References

- Robleda G, Roche-Campo F, Sanchez V, Gich I, Banos JE. Postoperative discomfort after abdominal surgery: An observational study. *J Perianesth Nurs.* 2015; 30:272-279.
- Hamalainen MM, Gebhart GF, Brennan TJ. Acute effect of an incision on mechanosensitive afferents in the plantar rat hindpaw. *J Neurophysiol.* 2002; 87:712-720.
- Yildirim V, Doganci S, Cinar S, Eskin MB, Ozkan G, Eksert S, Ince ME, Dogrul A. Acute high dose-fentanyl exposure produces hyperalgesia and tactile allodynia after coronary artery bypass surgery. *Eur Rev Med Pharmacol Sci.* 2014; 18:3425-3434.
- Wen YR, Lin CP, Tsai MD, Chen JY, Ma CC, Sun WZ, Wang CC. Combination of nerve blockade and intravenous alfentanil is better than single treatment in relieving postoperative pain. *J Formos Med Assoc.* 2012; 111:101-108.
- Yamada T, Hasegawa-Moriyama M, Kurimoto T, Saito T, Kuwaki T, Kanmura Y. Peripheral nerve block facilitates acute inflammatory responses induced by surgical incision in mice. *Reg Anesth Pain Med.* 2016; 41:593-600.
- Zhang YL, Xu JM, Zhou P, Zhong XL, Dai RP. Distinct activation of tumor necrosis factor- α and interleukin-6 in the spinal cord after surgical incision in rats. *Mol Med Rep.* 2012; 5:1423-1427.
- Wen YR, Suter MR, Ji RR, Yeh GC, Wu YS, Wang KC, Kohno T, Sun WZ, Wang CC. Activation of p38 mitogen-activated protein kinase in spinal microglia contributes to incision-induced mechanical allodynia. *Anesthesiology.* 2009; 110:155-165.
- Zahn PK, Brennan TJ. Primary and secondary hyperalgesia in a rat model for human postoperative pain. *Anesthesiology.* 1999; 90:863-872.
- Shu H, Hayashida M, Arita H, Huang W, Zhang H, An K, Wu G, Hanaoka K. Pentazocine-induced antinociception is mediated mainly by μ -opioid receptors and compromised by kappa-opioid receptors in mice. *J Pharmacol Exp Ther.* 2011; 338:579-587.
- Hargreaves K, Dubner R, Brown F, Flores C, Joris J. A new and sensitive method for measuring thermal nociception in cutaneous hyperalgesia. *Pain.* 1988; 32:77-88.
- Chang L, Ye F, Luo Q, Tao Y, Shu H. Increased hyperalgesia and proinflammatory cytokines in the spinal cord and dorsal root ganglion after surgery and/or fentanyl administration in rats. *Anesth Analg.* 2018; 126:289-297.
- Celerier E, Rivat C, Jun Y, Laulin JP, Larcher A, Reynier P, Simonnet G. Long-lasting hyperalgesia induced by fentanyl in rats: Preventive effect of ketamine. *Anesthesiology.* 2000; 92:465-472.
- Lee M, Silverman S, Hansen H, Patel V, Manchikanti L. A comprehensive review of opioid-induced hyperalgesia. *Pain Physician.* 2011; 14:145-161.
- Berta T, Liu T, Liu YC, Xu ZZ, Ji RR. Acute morphine activates satellite glial cells and up-regulates IL-1 beta in dorsal root ganglia in mice via matrix metalloproteinase-9. *Mol Pain.* 2012; 8:18.
- Mizukoshi K, Sasaki M, Izumi Y, Miura M, Watanabe M, Amaya F. Activation of p38 mitogen-activated protein kinase in the dorsal root ganglion contributes to pain hypersensitivity after plantar incision. *Neuroscience.* 2013; 234:77-87.
- Brennan TJ, Zahn PK, Pogatzki-Zahn EM. Mechanisms of incisional pain. *Anesthesiol Clin North America.* 2005; 23:1-20.
- Rivat C, Vera-Portocarrero LP, Ibrahim MM, Mata HP, Stagg NJ, De Felice M, Porreca F, Malan TP. Spinal NK-1 receptor-expressing neurons and descending pathways support fentanyl-induced pain hypersensitivity in a rat model of postoperative pain. *Eur J Neurosci.* 2009; 29:727-737.
- Celerier E, Gonzalez JR, Maldonado R, Cabanero D, Puig MM. Opioid-induced hyperalgesia in a murine model of postoperative pain: Role of nitric oxide generated from the inducible nitric oxide synthase. *Anesthesiology.* 2006; 104:546-555.
- Polomano RC, Mannes AJ, Clark US, Bennett GJ. A painful peripheral neuropathy in the rat produced by the chemotherapeutic drug, paclitaxel. *Pain.* 2001; 94:293-304.
- Barrevelde A, Witte J, Chahal H, Durieux ME, Strichartz G. Preventive analgesia by local anesthetics: The reduction of postoperative pain by peripheral nerve blocks and intravenous drugs. *Anesth Analg.* 2013; 116:1141-1161.
- Marques EM, Jones HE, Elvers KT, Pyke M, Blom AW, Beswick AD. Local anaesthetic infiltration for perioperative pain control in total hip and knee replacement: Systematic review and meta-analyses of short- and long-term effectiveness. *BMC Musculoskelet Disord.* 2014; 15:220.
- Rivat C, Bollag L, Richebe P. Mechanisms of regional anaesthesia protection against hyperalgesia and pain chronicization. *Curr Opin Anaesthesiol.* 2013; 26:621-625.
- Meleine M, Rivat C, Laboureyras E, Cahana A, Richebe P. Sciatic nerve block fails in preventing the development of late stress-induced hyperalgesia when high-dose fentanyl is administered perioperatively in rats. *Reg Anesth Pain Med.* 2012; 37:448-454.
- McMahon SB, Cafferty WB, Marchand F. Immune and glial cell factors as pain mediators and modulators. *Exp Neurol.* 2005; 192:444-462.
- Basbaum AI, Bautista DM, Scherrer G, Julius D. Cellular and molecular mechanisms of pain. *Cell.* 2009; 139:267-284.
- Reagan-Shaw S, Nihal M, Ahmad N. Dose translation from animal to human studies revisited. *FASEB J.* 2008; 22:659-661.

(Received January 30, 2018; Revised March 11, 2018; Accepted March 27, 2018)

Infiltration characteristics and influencing factors of retroperitoneal liposarcoma: Novel evidence for extended surgery and a tumor grading system

Zhen Wang¹, Jianhui Wu¹, Ang Lv¹, Chengpeng Li¹, Zhongwu Li², Min Zhao², Chunyi Hao^{1,*}

¹ Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education/Beijing), Department of Hepato-Pancreato-Biliary Surgery, Peking University Cancer Hospital and Institute, Beijing, China;

² Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education/Beijing), Department of Pathology, Peking University Cancer Hospital and Institute, Beijing, China.

Summary

This study sought to evaluate the infiltration tendency of retroperitoneal liposarcoma (RPLS) from a new pathological angle by exploring the infiltration characteristics, which could provide helpful information to facilitate surgical decision-making and prognosis prediction. Concurrently, we aim to identify significant indicators of infiltration. A total of 61 consecutive patients with RPLS at our institution were retrospectively analyzed. All patients received extended surgery. The tumor infiltration characteristics and influencing factors were studied based on the pathological diagnosis. Univariate and multivariate analyses of organ infiltration (OI) and surrounding fat infiltration (SFI) were performed. OI was found in 95 (28.5%) resected organs from 39 (60.7%) patients, and SFI was found in 119 (35.7%) resected organs from 47 (77%) patients. The tumor infiltrated the serosal layer of 13 organs (13/37, 35.1%), the muscularis layer of 18 organs (18/37, 48.6%) and the submucosa of 6 organs (6/37, 16.2%). The percentage of lipoblasts and the rates of necrosis and mitosis were all significantly higher in high-grade tumors (dedifferentiated, round cell, and pleomorphic). A high lipoblast percentage ($\geq 20\%$) was the only independent risk factor for OI. A recurrent tumor and a high-grade tumor were independent risk factors for SFI. In conclusion, RPLS has a high infiltration tendency, such that it frequently infiltrates organs and surrounding fat tissue. Therefore, extended resection of the tumor and the adjacent organs is recommended. The percentage of lipoblasts was associated with the tumor grade and infiltration characteristics; thus, lipoblast percentage may become a new grading factor for RPLS.

Keywords: Retroperitoneal sarcoma, liposarcoma, infiltration

1. Introduction

Liposarcoma is the most common histotype of retroperitoneal sarcoma (RPS) (1). Local recurrence of RPS was reported to be the leading cause of death

in patients with this type of tumor (1,2). Pathological evaluation of the involved surgical margins is a major predictor of local tumor recurrence (1,3,4). To reduce the local recurrence, achieving negative surgical margins by extended resection of the tumor and the adjacent organs is recommended (5-7). The above recommendations were mostly based on analyses of clinical outcomes, such as local recurrence, metastasis and survival. The necessity of extended surgery can be directly evaluated by tumor infiltration based on the postoperative pathology. However, to the best of our knowledge, there are few studies offering details of retroperitoneal liposarcoma (RPLS) infiltration. RPS is reported to have a high rate of visceral infiltration (8); however, not all RPSs have similar infiltration characteristics (8,9).

Released online in J-STAGE as advance publication April 15, 2018.

*Address correspondence to:

Dr. Chunyi Hao, Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education/Beijing), Department of Hepato-Pancreato-Biliary Surgery, Peking University Cancer Hospital and Institute, #52 Fucheng Road, Haidian District, Beijing 100142, China.
E-mail: haochunyi@bjmu.edu.cn

Therefore, we conducted a single-center study analyzing a large series of patients with RPLS who underwent extended surgery to explore the infiltration characteristics and influencing factors of RPLS, which could provide helpful information to facilitate surgical decision-making and prognosis prediction.

Despite the widespread application of some grading systems in the diagnosis and management of sarcomas, there is no accurate grading system for liposarcoma (10). Studying the clinico-pathological characteristics of RPLS, especially pathological details after extended resection, can help us establish a new grading rationale, which will provide more information for developing a more reliable grading system.

2. Materials and Methods

2.1. Patients

We retrospectively analyzed a total of 61 consecutive patients with RPLS from the Peking University Cancer Hospital Sarcoma Center between March 2015 and March 2017. The median follow-up time was 11 months (interquartile range [IQR], 7-20 months). All patients included in the study underwent surgery with curative intent. Pancreaticoduodenectomy was conducted when the tumor infiltrated the pancreatic head and greater part of the duodenum. If necessary, great vessels (such as the inferior vena cava, aorta and iliac vessels) were removed and replaced with polytetrafluoroethylene grafts. If the diaphragm and/or pericardium showed tumor invasion, partial diaphragm and/or pericardium resection and repair were performed. Our surgical policy was to remove the tumor with adjacent organs *en bloc*, which is more aggressive than most surgeries described in past studies (5,7,8).

2.2. Pathological diagnosis

All tumors were delivered to the pathology receiving room after the operation. Overall tumor size was defined as the sum of the perpendicular maximum diameters of the primary tumors as reported at the time of initial surgical resection, and tumors were grouped into three overall size categories: < 15 cm, 15-30 cm, and \geq 30 cm. The specimen was orientated by the surgeon. All margins were perpendicularly sampled, with two or more sections taken from all margins. Additional sections were taken from the closest margin. Serial sampling of all resected organs and surrounding fat was performed, and the tissue between the tumor and organ were sampled every 2 cm.

Two sarcoma pathologists independently conducted the pathological diagnosis. The rate of necrosis, percentage of lipoblasts, and number of mitotic events per high-powered field were all recorded. The results of the serially sampled specimens are presented as their

mean lipoblast percentage.

2.3. Definitions

The liposarcomas were subdivided into 5 pathological subtypes: well-differentiated, dedifferentiated, myxoid (< 5% round cell component), round cell (\geq 5% round cell), or pleomorphic. Additionally, the liposarcomas were divided into a high-grade group (dedifferentiated, round cell, and pleomorphic) and a low-grade group (well-differentiated and myxoid) (11).

Surrounding fat was defined as the fat tissue within 1 mm from the organ surface. The infiltration pattern was classified as organ infiltration (OI) and surrounding fat infiltration (SFI). OI was defined if infiltration of the bowel or parenchyma of the solid organ was observed. SFI was defined if only infiltration of the surrounding fat was observed, without OI.

Lipoblasts have hyperchromatic, indented or sharply scalloped nuclei and a lipid-rich mono- or multivacuolated cytoplasm (10). Intracytoplasmic vacuoles appear sharply margined, which is distinct from the ill-defined vacuoles of pseudolipoblasts that contain mucin. The nucleus of a lipoblast may be pushed aside by a single large lipid vacuole, resulting in a signet-ring configuration, or it may remain centrally located with small indentations caused by multiple, small vacuoles, resulting in an appearance similar to that of sebaceous cells or adrenal spongocytes. In addition to the above typical lipoblast characteristics, atypical or giant forms having large, irregular or multiple nuclei may be present (12).

Surgical resection was described as macroscopically complete (R0 or R1) or incomplete (R2) (3).

2.4. Postoperative complications and follow-up

Postoperative complications were graded according to the Clavien-Dindo classification (13). The patients were prospectively followed by clinical examination, chest X-ray, and abdominopelvic computed tomography (CT) or magnetic resonance imaging (MRI) every three months for the first two years, every six months for the subsequent three years, and yearly thereafter.

2.5. Data analysis

Data are presented as median and range, or number and percentage, where appropriate. Variables with a p-value \leq 0.1 were considered in the multivariate models. Multivariable analysis of OI and SFI was carried out using logistic regression models. For patients who underwent macroscopically complete (R0 or R1) surgical resection, we analyzed the local disease-free survival (DFS) from the date of operation to the date of the last follow-up. Statistical analyses were performed using SPSS version 24.0 (Armonk, NY; IBM Corp) and

R version 3.4.0 (<http://www.r-project.org>). *P*-values less than 0.05 were considered statistically significant.

3. Results

3.1. Patient characteristics

Sixty-one patients were enrolled in our study. Thirty-two (52.5%) patients received primary treatment, and the remaining patients received operation after tumor recurrence. All the operations performed before recurrence were simple excisions of the tumor (along with adjacent organs only if visibly infiltrated). Complete macroscopic (R0 or R1) resection was achieved in 55 cases (90.2%). The most common histological type was dedifferentiated liposarcomas (28/61, 45.9%), and there were 17 (27.9%) well-differentiated liposarcomas, 3 (4.9%) myxoid liposarcomas, 4 (6.6%) round cell liposarcomas and 9 (14.8%) pleomorphic liposarcomas. The clinical-pathological characteristics of the entire series are displayed in Table 1.

3.2. Overview of the operation and infiltration

Three hundred and thirty-three organs were resected (5.5/patient). Colectomy was the most frequent

Table 1. Patients and characteristics

Characteristics	N (%)
Age [years; median (IQR)]	56 (48-63)
Sex	
Male	33 (54.1)
Female	28 (45.9)
Presentation	
Primary	32 (52.5)
Recurrent	29 (47.5)
Tumor site	
Left side	13 (21.3)
Right side	17 (27.9)
More than one area	31 (50.8)
Tumor number	
Single	46 (75.4)
Multiple	15 (24.6)
Tumor size	
< 15 cm	11 (18.0)
15-30 cm	31 (50.8)
≥ 30 cm	19 (31.1)
Number of organs resected [median (IQR)]	5 (2-8)
Histological subtype	
Well-differentiated	17 (27.9)
Dedifferentiated	28 (45.9)
Myxoid	3 (4.9)
Round cell	4 (6.6)
Pleomorphic	9 (14.8)
Resection margins	
Macroscopically complete	55 (90.2)
Macroscopically incomplete	6 (9.8)
Grade	
Low-grade	20 (32.8)
High-grade	41 (67.2)

IQR, interquartile range.

procedure performed (50/61, 82.0%). Forty-nine (80.3%) patients had at least one organ and/or surrounding fat tissue infiltrated by the tumor. OI was found in 95 resected organs (95/333, 28.5%) of 39 (39/61, 60.7%) patients, and when the organ was infiltrated by the tumor, most of the fat surrounding the organ was also found to be infiltrated. SFI was found in 119 resected organs (119/333, 35.7%) of 47 (47/61, 77%) patients. Mesenteric fat was infiltrated in 13 patients (13/61, 21.3%) who underwent colectomy and/or resection of the small intestine. The details of OI and SFI of resected organs (> 10 patients) are displayed in Figure 1.

3.3. Infiltration of the hollow viscera

For the hollow viscera (including the colon, small intestine and stomach), 37 organs (37/95, 38.9%) were infiltrated by the tumor. Of these, the tumor infiltrated the serosal layer of 13 organs (13/37, 35.1%), the muscularis of 8 organs (18/37, 48.6%) and the submucosa of 6 organs (6/37, 16.2%) (Figure 2). The

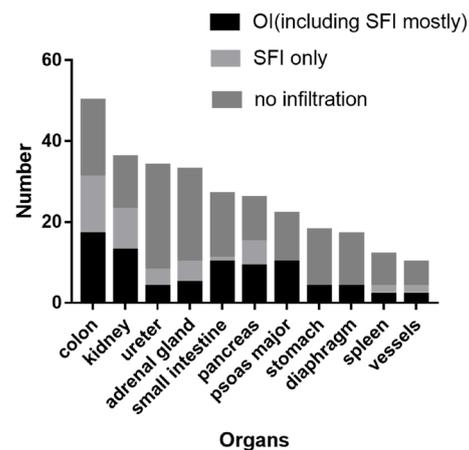


Figure 1. The details of OI and SFI of resected organs (> 10 patients). OI (including SFI in most cases): organ infiltration. When the organ is infiltrated by the tumor, most of the fat surrounding the organ is simultaneously infiltrated. SFI only: only surrounding fat infiltration was observed, without OI.

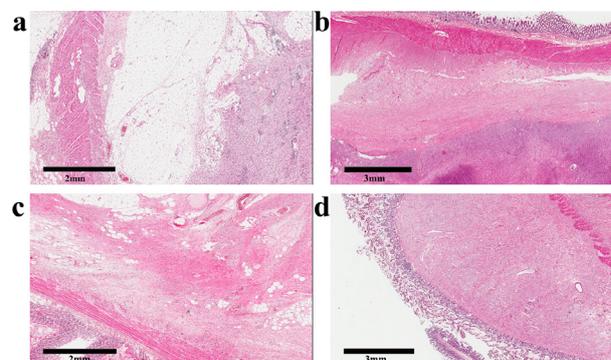


Figure 2. Infiltration of the hollow viscera. (a) Liposarcoma infiltration of the surrounding fat tissue of the colon; (b) Liposarcoma infiltration of the serosal layer of the colon; (c) Liposarcoma infiltration of the muscularis of the colon; (d) Liposarcoma infiltration of the submucosa of the duodenum.

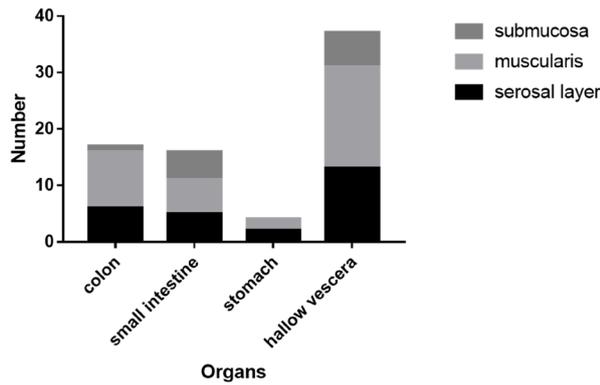


Figure 3. Infiltration details of the hollow viscera.

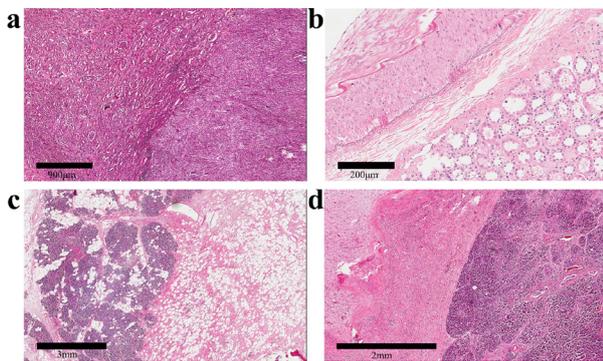


Figure 4. Infiltration of the kidney and pancreas. (a) Liposarcoma infiltration observed in the renal parenchyma; (b) Liposarcoma infiltration observed in only the perirenal fat sac, without the renal parenchyma being infiltrated; (c) Liposarcoma infiltration observed in the pancreatic parenchyma; (d) Liposarcoma infiltration observed in only the surrounding fat tissue of the pancreas, without the pancreatic parenchyma being infiltrated.

colon was the most frequently infiltrated organ (17/50, 34.0%). Of the colon samples, the tumor infiltrated the serosal layer in 6 patients (6/17, 35.3%), the muscularis in 10 patients (10/17, 58.8%) and the submucosa in 1 patient (1/17, 5.9%). Fourteen (14/50, 28.0%) patients showed infiltration of the surrounding fat tissue without colon infiltration (Figure 2a). Infiltration details of the hollow viscera are displayed in Figure 3.

3.4. Infiltration of the kidney and pancreas

Of the patients who underwent nephrectomy, 13 (13/36, 36.1%) were found to have tumor infiltration of the renal parenchyma and 10 (10/36, 27.8%) were found to have tumor infiltration in only the perirenal fat sac (Figure 4a and 4b). Of the patients who underwent a pancreas resection, 9 (9/26, 34.6%) were found to have tumor infiltration of the pancreatic parenchyma and 6 (6/26, 23.1%) showed tumor infiltration in only the surrounding fat tissue (Figure 4c and Figure 4d).

3.5. Necrosis, lipoblast percentage, mitosis and grade

Twenty-six (26/61, 42.6%) patients showed tumor

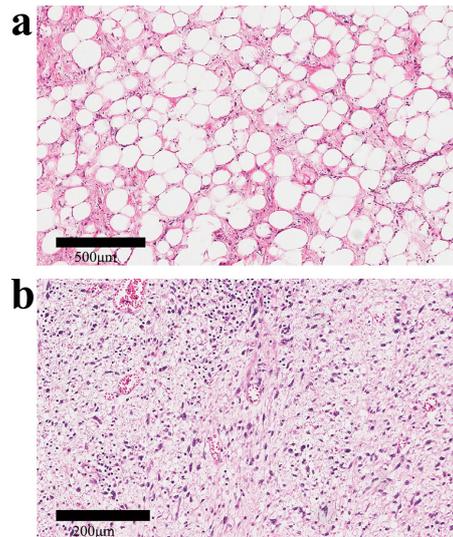


Figure 5. Lipoblast percentage. (a) Low lipoblast percentage ($\leq 20\%$); (b) High lipoblast percentage ($\geq 20\%$).

necrosis. Necrosis was found in 3 patients (3/20, 15%) with low-grade tumors and 23 patients (23/41, 56.1%) with high-grade tumors ($p = 0.002$). The median percentage of lipoblasts was 10% (IQR 5%-32.5%). In patients with low-grade tumors, the median percentage of lipoblasts was 5%, and 80.0% of patients had percentages no higher than 5% (Figure 5a). However, the median percentage of lipoblasts was 20% (IQR 5%-40%) in patients with high-grade tumors (Figure 5b). There was a statistically significant difference between the two groups ($p < 0.001$). Mitosis was observed in the tumors of 29 patients (29/61, 47.5%), and the median number of mitotic events per high-powered field was 3 (IQR 1.5-5.5) in these patients. Mitosis was found in the tumors of 3 patients (3/20, 15%) with low-grade tumors and 26 patients (26/41, 63.4%) with high-grade tumors ($p < 0.001$).

3.6. Risk factors of OI and SFI

The tumor site ($p = 0.037$), a high tumor grade ($p = 0.002$), necrosis ($p = 0.022$) and a high lipoblast percentage ($\geq 20\%$) ($p = 0.002$) were all significantly associated with OI in the univariate analysis. In the multivariate analysis, these associations were confirmed for the high lipoblast percentage ($\geq 20\%$) (OR: 11.67; 95% CI: 2.39-56.87; $p = 0.002$).

In the univariate analysis, recurrent tumor ($p = 0.034$), a high tumor grade ($p = 0.001$) and a high lipoblast percentage ($\geq 20\%$) ($p = 0.024$) were all significantly associated with SFI. Multivariate analysis showed that recurrent tumors (OR: 6.18; 95% CI: 1.24-30.75; $p = 0.026$) and high-grade tumors (OR: 11.62; 95% CI: 2.61-51.64; $p = 0.001$) were independent risk factors for SFI. The univariate and multivariate analyses of risk factors of OI and SFI are shown in Table 2.

Table 2. Univariate and multivariate analyses of OI and SFI

Characteristics	OI			SFI		
	Univariate analysis <i>p</i> -value	Multivariate analysis OR (95% CI)	<i>p</i> -value	Univariate analysis <i>p</i> -value	Multivariate analysis OR (95% CI)	<i>p</i> -value
Age	0.143			0.232		
Sex	0.557			0.340		
Male						
Female						
Presentation	0.437			0.034		0.026
Primary					1	
Recurrent					6.18 (1.24-30.75)	
Tumor site						
Left side	0.042	1	0.371	0.358		
Right side	0.583	0.49 (0.06-4.18)	0.516	0.271		
More than one area	0.037	0.37 (0.06-2.43)	0.172	0.153		
Tumor number	0.115			0.276		
Single						
Multiple						
Tumor size						
< 15 cm	0.267			0.691		
15-30 cm	0.156			0.758		
≥ 30 cm	0.539			0.421		
Grade	0.002		0.155	0.001		0.001
Low-grade		1			1	
High-grade		2.89 (0.67-12.50)			11.62 (2.61-51.64)	
Necrosis	0.022		0.486	0.232		
No		1				
Yes		1.79 (0.35-9.10)				
Lipoblast percentage	0.002		0.002	0.024		0.089
< 20%		1			1	
≥ 20%		11.67 (2.39-56.87)			7.53 (0.74-76.87)	
Mitosis	0.069		0.735	0.114		
No		1				
Yes		0.76 (1.50-3.75)				

IQR, interquartile range; OI, organ infiltration; SFI, surrounding fat infiltration.

3.7. Postoperative morbidity

Grade 3-4 postoperative morbidity occurred in 13 (13/61, 21.3%) patients. Grade 3a, 3b and 4 occurred in 6 (6/61, 9.8%), 3 (3/61, 4.9%), and 4 patients (4/61, 6.6%), respectively. Four patients had abdominal infections. Three patients had grade B postoperative pancreatic fistulas (POPF). Four patients had postoperative hemorrhages: 2 patients were due to gastroduodenal artery bleeding; 1 patient was due to gastrointestinal anastomosis bleeding and 1 patient was due to rupture of the iliac artery and artificial vessel anastomosis. One patient had necrosis of the left lower extremity, and thrombosis of the inferior vena cava occurred in 1 patient. Surgical reintervention was necessary in 7 patients (7/61, 11.5%). Three patients underwent surgery for abdominal infection, 2 for gastroduodenal arterial bleeding, 1 for rupture of the iliac artery and artificial vessel anastomosis, and 1 for necrosis of the left lower extremity. Three patients (3/61, 4.9%) died of postoperative complications within 90 days of surgery. Of these, 1 patient underwent reoperation for abdominal infection. The 2 remaining deaths were due to myocardial infarction and renal failure, respectively.

3.8. Follow-up

Of the 55 patients who underwent macroscopically complete (R0 or R1) resection, 3 patients (3/55, 5.6%) experienced local recurrence, and 1 patient experienced death after recurrence. The 1-year DFS was 91.2% (95% CI: 83.2%-100.0%) in patients with macroscopically complete resection. Of the 6 patients who underwent R2 resection during the operation, 3 patients died of tumor progression. The 1-year overall survival (OS) was 91.2% (95% CI: 84.1%-99.9%) in all the patients who received extended surgery.

The patient's ability to perform general daily activities was scored according to a modified Barthel Index (mBI) (14). Fifty-seven patients (93.4%) scored 100, and the remaining patients scored higher than 70.

4. Discussion

In this study, 61 consecutive patients with RPLS were surgically treated at our center. Reviewing the pathological results, we found that the infiltration of RPLS had several characteristics. As RPLS tends to infiltrate adjacent organs, the tumor infiltrated at least one organ and/or the surrounding fat tissue in

most patients (49/61, 80.3%) evaluated in our study. In addition, the RPLS infiltrated the deep layer of the cavity organs. The above results indicate that RPLS has a strong ability to infiltrate surrounding areas. This conclusion is of great importance because surgeons too often believe that RPLS is a mass that only presses against but does not infiltrate the adjacent organs; consequently, patients undergo simple resection of the tumor itself and may be at increased risk of more frequent tumor recurrence. Our notion supports a policy of extended surgery beyond visible infiltrations.

In our study, extended resection of the tumor and adjacent kidney was performed based on the preoperative imaging features and intra-operative exploration. However, the postoperative pathological diagnosis showed only 36.1% of the patients who underwent nephrectomy had renal parenchyma infiltration, and 27.8% were found to have tumor infiltration in only the perirenal fat sac. The above infiltration characteristics may suggest the reasonability to preserve these uninvaded organs. However, there are no current techniques to evaluate RPLS infiltration accurately before or during the operation. Preserving kidneys without reliable evaluation of parenchymal infiltration may result to the increase probability of recurrence. Moreover, previous studies have indicated that renal resection during the primary surgery provided better DFS (15,16). Furthermore, a published study with more than four years of follow-up revealed that patients who underwent renal resection did not require treatment for renal insufficiency, even among patients who received postoperative chemotherapy (17). Additionally, during kidney-preserving surgery, exploration is necessary to determine whether the tumor has invaded the renal parenchyma, which may lead to rupture of the RPLS and increase the probability of recurrence and peritoneal metastasis (18). Therefore, nephrectomy is recommended for RPLS that occurs adjacent to the kidney. The detailed imaging features of patients with renal parenchymal infiltration needs to be studied in the future, which can help with determination of renal parenchymal infiltration before operation.

In this study, we showed for the first time that a considerable proportion of resected pancreases had only SFI and were free of parenchymal infiltration, even though extended resection of the tumor and adjacent pancreas was performed based on the preoperative imaging features and intra-operative exploration. Currently, no technique can accurately assess parenchymal infiltration before or during surgery. Preserving pancreas without reliable evaluation of parenchymal infiltration results to the increase probability of positive surgical margin and local recurrence. Furthermore, it was reported that the clinical outcomes of individuals with recurrent tumors were worse than primary tumors (19-21). In addition, although POPF accounted for 23.1% of all Grade 3 and 4

complications in the present study, drainage was effective in those cases. Furthermore, no long-term morbidity related to pancreas resection occurred during the follow-up. Therefore, organ preservation will be possible only if preoperative or intra-operative techniques are developed to accurately assess adjacent organ involvement.

Grade 3-4 postoperative complications occurred in 13 patients (21.3%), and surgical reintervention was necessary in 7 patients (11.5%). These results are comparable to those of previous studies of extended resection for RPS, which revealed severe postoperative complication rates of 16.4-30% and surgical reintervention rates of 12-14% (5,22,23). Bonvalot *et al.* (5) and MacNeill *et al.* (23) reported that removing organs can increase postoperative complications. However, these previous studies also showed that postsurgical morbidity did not affect the oncological outcome (5,23). Considering the strong infiltrative ability of RPLS mentioned above, preserving the organs and removing the invaded surrounding fat necessitates only the surgeon's careful consideration on the recurrent risks and benefits. Our data also showed that the perioperative mortality, rate of postoperative morbidity and rate of reoperation after extended surgery were acceptable. Currently, extended resection including the tumor and adjacent organs should be suggested to achieve radical treatment.

Furthermore, our study found that recurrent and high-grade tumors can significantly increase the incidence of SFI. This association may be the cause of the high local DFS of patients with recurrent and high-grade tumors observed in previous studies (1,24,25). Previous studies have shown that patients with high-grade sarcomas do not significantly benefit from extended resection (7,26) and that distant metastasis was a limiting factor. Lehnert T *et al.* (21) reported that consequent reoperation led to satisfactory long-term survival rates after resection of recurrent tumors. However, Gronchi *et al.* (19) reported that a second surgery is of limited benefit for individuals with recurrent tumors. Park JO *et al.* (20) reported that only patients with local recurrence growth rates slower than 0.9 cm per month were associated with improved survival after aggressive resection of the local recurrence. There is no consensus on the treatment of recurrent tumors. If an operation is conducted for recurrent and high-grade tumors, a more extended resection including retroperitoneal fat and tumor adjacent organs is the recommended approach to reduce local DFS because of their higher rate of SFI.

Lipoblasts are conceptually an immature or precursor form of adipocytes and are essentially identified in neoplastic conditions (12). Traditionally, great emphasis has been placed on the identification of lipoblasts as a diagnostic marker for liposarcoma (10). However, as far as we know, the relation between the lipoblast count and the tumor grade or degree of

infiltration has not been explored in previous studies. In our study, quantitative analysis of lipoblasts was carried out, and interestingly, we found that an increase in lipoblast percentage can significantly increase the incidence of OI (OR: 11.67; 95% CI: 2.39-56.87; $p = 0.002$). Moreover, there was also a tendency to increase the occurrence of SFI (OR: 7.53; 95% CI: 0.74-76.87; $p = 0.089$). Therefore, lipoblast percentage can be considered a new indicator to predict the ability of infiltration.

Our results show that the rate of necrosis was significantly increased in high-grade tumors. The grading system published by the National Cancer Institute (NCI) (27) and the French Federation of Cancer Centers Sarcoma Group (FNCLCC) (28) both consider necrosis and mitosis as indicators of oncological outcomes. However, the percentage of lipoblasts has never been considered in previous studies (27,29-31). Some studies have shown that RPLS infiltration of adjacent organs is associated with a poor prognosis (25), and an increase in the percentage of lipoblasts can significantly increase the infiltration of RPLS. Thus, the lipoblast percentage can also be considered a grading factor for RPLS. As this study was limited by its short follow-up time, the relationship between lipoblast percentage and oncological outcome cannot be verified for the time being.

To improve the prognosis of patients with RPLS, extended resection of the tumor and the adjacent organs is recommended (5-7). In the present study, the 1-year DFS rate was 91.2% in patients with macroscopically complete resection. Of the 6 patients who underwent R2 resection during the operation, 3 patients died of tumor progression. Although the follow-up period is short, the follow-up results to date are encouraging. Furthermore, during the follow-up period, all living patients had a high quality of life. However, a larger sample size and longer follow-up time will help further validate the efficacy of extended surgery in our center.

In conclusion, RPLS has a high infiltrative tendency, such that it frequently infiltrates organs and surrounding fat. Meanwhile, the perioperative mortality, rate of postoperative morbidity and rate of reoperation after extended surgery were deemed acceptable. Therefore, extended resection of the tumor and the adjacent organs should be recommended. Moreover, the percentage of lipoblasts was found to be associated with the tumor grade and infiltration characteristics. Thus, the lipoblast percentage may become a new grading factor for RPLS.

Acknowledgements

We thank all faculty members who assisted us in this study. This work was supported by Beijing Municipal Administration of Hospitals Clinical Medicine Development Special Funding Support (approval No.: XMLX201708), the Capital Health Research and

Development Special Funds (approval No.: 2016-2-2151), and National Natural Science Funding (approval No.: 31770836).

References

- Lewis JJ, Leung D, Woodruff JM, Brennan MF. Retroperitoneal soft-tissue sarcoma: Analysis of 500 patients treated and followed at a single institution. *Ann Surg.* 1998; 228:355-365.
- Linehan DC, Lewis JJ, Leung D, Brennan MF. Influence of biologic factors and anatomic site in completely resected liposarcoma. *J Clin Oncol.* 2000; 18:1637-1643.
- Anaya DA, Lev DC, Pollock RE. The role of surgical margin status in retroperitoneal sarcoma. *J Surg Oncol.* 2008; 98:607-610.
- Mankin HJ, Mankin KP, Harmon DC. Liposarcoma: A soft tissue tumor with many presentations. *Musculoskelet Surg.* 2014; 98:171-177.
- Bonvalot S, Miceli R, Berselli M, Causeret S, Colombo C, Mariani L, Bouzaïene H, Le Pechoux C, Casali PG, Le Cesne A, Fiore M, Gronchi A. Aggressive surgery in retroperitoneal soft tissue sarcoma carried out at high-volume centers is safe and is associated with improved local control. *Ann Surg Oncol.* 2010; 17:1507-1514.
- Porter GA, Baxter NN, Pisters PW. Retroperitoneal sarcoma: A population-based analysis of epidemiology, surgery, and radiotherapy. *Cancer.* 2006; 106:1610-1616.
- Gronchi A, Miceli R, Colombo C, Stacchiotti S, Collini P, Mariani L, Sangalli C, Radaelli S, Sanfilippo R, Fiore M, Casali PG. Frontline extended surgery is associated with improved survival in retroperitoneal low- to intermediate-grade soft tissue sarcomas. *Ann Oncol.* 2012; 23:1067-1073.
- Mussi C, Colombo P, Bertuzzi A, Coladonato M, Bagnoli P, Secondino S, Navarra P, Morengi E, Santoro A, Quagliuolo V. Retroperitoneal sarcoma: Is it time to change the surgical policy? *Ann Surg Oncol.* 2011; 18:2136-2142.
- Gronchi A, Pollock RE. Quality of local treatment or biology of the tumor: Which are the trump cards for loco-regional control of retroperitoneal sarcoma? *Ann Surg Oncol.* 2013; 20:2111-2113.
- Enzinger FM, Weiss SW, Folpe AL, Weiss SW, Goldblum JR. *Enzinger and Weiss's soft tissue tumors.* Saunders/Elsevier, Philadelphia, PA, USA, 2014; pp. 1-10.
- Tan MC, Brennan MF, Kuk D, Agaram NP, Antonescu CR, Qin LX, Moraco N, Crago AM, Singer S. Histology-based classification predicts pattern of recurrence and improves risk stratification in primary retroperitoneal sarcoma. *Ann Surg.* 2016; 263:593-600.
- Hisaoka M. Lipoblast: Morphologic features and diagnostic value. *J UOEH.* 2014; 36:115-121.
- Dindo D, Demartines N, Clavien PA. Classification of surgical complications: A new proposal with evaluation in a cohort of 6336 patients and results of a survey. *Ann Surg.* 2004; 240:205-213.
- Loewen SC, Anderson BA. Reliability of the modified motor assessment scale and the barthel index. *Phys Ther.* 1988; 68:1077-1081.
- Rhu J, Cho CW, Lee KW, Park H, Park JB, Choi Y-L, Kim SJ. Radical nephrectomy for primary retroperitoneal liposarcoma near the kidney has a beneficial effect on

- disease-free survival. *World J Surg.* 2018; 42:254-262.
16. Singer S, Antonescu CR, Riedel E, Brennan MF. Histologic subtype and margin of resection predict pattern of recurrence and survival for retroperitoneal liposarcoma. *Ann Surg.* 2003; 238:358-370; discussion 370-351.
 17. Hull MA, Niemierko A, Haynes AB, Jacobson A, Chen Y-L, DeLaney TF, Mullen JT. Post-operative renal function following nephrectomy as part of en bloc resection of retroperitoneal sarcoma (RPS). *J Surg Oncol.* 2015; 112:98-102.
 18. Rafat C, Zinzindohoue F, Hernigou A, Hignette C, Favier J, Tenenbaum F, Gimenez-Roqueplo AP, Plouin PF, Amar L. Peritoneal implantation of pheochromocytoma following tumor capsule rupture during surgery. *J Clin Endocrinol Metab.* 2014; 99:E2681-E2685.
 19. Gronchi A, Miceli R, Allard MA, *et al.* Personalizing the approach to retroperitoneal soft tissue sarcoma: Histology-specific patterns of failure and postrelapse outcome after primary extended resection. *Ann Surg Oncol.* 2015; 22:1447-1454.
 20. Park JO, Qin LX, Prete FP, Antonescu C, Brennan MF, Singer S. Predicting outcome by growth rate of locally recurrent retroperitoneal liposarcoma: The one centimeter per month rule. *Ann Surg.* 2009; 250:977-982.
 21. Lehnert T, Cardona S, Hinz U, Willeke F, Mechtersheimer G, Treiber M, Herfarth C, Buechler MW, Schwarzbach MH. Primary and locally recurrent retroperitoneal soft-tissue sarcoma: Local control and survival. *Eur J Surg Oncol.* 2009; 35:986-993.
 22. Pasquali S, Vohra R, Tsimopoulou I, Vijayan D, Gourevitch D, Desai A. Outcomes following extended surgery for retroperitoneal sarcomas: Results from a UK referral centre. *Ann Surg Oncol.* 2015; 22:3550-3556.
 23. MacNeill AJ, Gronchi A, Miceli R, *et al.* Postoperative morbidity after radical resection of primary retroperitoneal sarcoma: A report from the transatlantic RPS working group. *Ann Surg.* 2018; 267:959-964.
 24. Bonvalot S, Rivoire M, Castaing M, Stoeckle E, Le Cesne A, Blay JY, Laplanche A. Primary retroperitoneal sarcomas: A multivariate analysis of surgical factors associated with local control. *J Clin Oncol.* 2009; 27:31-37.
 25. Toulmonde M, Bonvalot S, Ray-Coquard I, *et al.* Retroperitoneal sarcomas: Patterns of care in advanced stages, prognostic factors and focus on main histological subtypes: A multicenter analysis of the French Sarcoma Group. *Ann Oncol.* 2014; 25:730-734.
 26. Gronchi A, Vullo SL, Fiore M, Mussi C, Stacchiotti S, Collini P, Lozza L, Pennacchioli E, Mariani L, Casali PG. Aggressive surgical policies in a retrospectively reviewed single-institution case series of retroperitoneal soft tissue sarcoma patients. *J Clin Oncol.* 2009; 27:24-30.
 27. Costa J, Wesley RA, Glatstein E, Rosenberg SA. The grading of soft tissue sarcomas. Results of a clinicohistopathologic correlation in a series of 163 cases. *Cancer.* 1984; 53:530-541.
 28. Trojani M, Contesso G, Coindre JM, Rouesse J, Bui NB, de Mascarel A, Goussot JF, David M, Bonichon F, Lagarde C. Soft-tissue sarcomas of adults; study of pathological prognostic variables and definition of a histopathological grading system. *Int J Cancer.* 1984; 33:37-42.
 29. van Unnik JA, Coindre JM, Contesso C, Albus-Lutter CE, Schiodt T, Sylvester R, Thomas D, Bramwell V, Mouridsen HT. Grading of soft tissue sarcomas: Experience of the EORTC Soft Tissue and Bone Sarcoma Group. *Eur J Cancer.* 1993; 29A:2089-2093.
 30. Jensen OM, Høgh J, Ostgaard SE, Nordentoft AM, Sneppen O. Histopathological grading of soft tissue tumours. Prognostic significance in a prospective study of 278 consecutive cases. *J Pathol.* 1991; 163:19-24.
 31. Hashimoto H, Daimaru Y, Takeshita S, Tsuneyoshi M, Enjoji M. Prognostic significance of histologic parameters of soft tissue sarcomas. *Cancer.* 1992; 70:2816-2822.

(Received January 31, 2018; Revised April 6, 2018; Accepted April 10, 2018)

A reliable grading system for prediction of hematoma expansion in intracerebral hemorrhage in the basal ganglia

Yongwei Huang¹, Qiang Zhang^{2,*}, Mingfei Yang^{2,*}

¹ Graduate School, Qinghai University, Xining, Qinghai, China;

² Department of Neurosurgery, Qinghai Provincial People's Hospital, Xining, Qinghai, China.

Summary

Hematoma expansion (HE) is an independent predictor of poor outcome and secondary neurological deterioration in intracerebral hemorrhage (ICH) and is associated with high morbidity and mortality. Noncontrast computed tomography (NCCT) may identify the sites of active extravasation. Therefore, we have attempted to (1) devise a reliable and easy-to-use prediction score to predict the risk of HE in ICH and (2) validate the accuracy of this grading system and perform an independent analysis of HE predictors. We included patients in whom an intracerebral hemorrhage (ICH) occurred in the basal ganglia between Jan. 2015 and Jan. 2018. These patients had undergone a baseline CT scan at Qinghai Provincial People's Hospital within 24 hours after the onset of ICH symptoms. Two observers independently assessed the presence of the island sign, blend sign, or swirl sign on an NCCT scan during patient selection. Patients underwent a baseline NCCT scan and 24-hour NCCT follow-up for analysis of HE. The accuracy of this grading system was assessed. Independent predictors of HE were identified using multivariable regression. Of 266 patients with ICH, 61 (22.93%) presented with the island sign, 63 (23.68%) presented with the blend sign, and 50 (18.80%) presented with the swirl sign. The overall incidence of HE was 37.22% (99/266). Of 125 patients (46.99%) who underwent a baseline CT scan within 6 hours of onset, 141 (53.01%) underwent a scan in 6-24 hours. Multivariable logistic regression analysis identified the hematoma volume (OR, 0.974; $P = 0.042$), intraventricular hemorrhage (IVH) extension (OR, 3.225; $P = 0.003$), time from onset to the baseline CT scan (OR, 0.986; $P < 0.001$), and anticoagulant use or an international normalized ratio (INR) > 1.5 (OR, 3.362; $P = 0.006$) as closely associated with HE. In conclusion, the grading system demonstrated reliable accuracy at predicting HE. The grading system demonstrated acceptable accuracy in an independent single-institution study. The role of the grading system in predicting HE and poor outcome in patients with ICH is significant. NCCT imaging markers may serve as key markers for HE prediction.

Keywords: Hematoma expansion, intracerebral hemorrhage, noncontrast computed tomography, grading system, basal ganglia

1. Introduction

Intracerebral hemorrhage (ICH) is a common disorder of the cerebral small vessels. ICH is a dynamic disorder, with up to 1/3 of patients suffering continued bleeding

after initial onset (1). ICH accounts for about 10% to 30% of all strokes worldwide (2) and it has a poor prognosis with a morbidity and mortality approaching 50% at 30 days (3).

The mechanism of hematoma expansion (HE) in ICH is not clear at this time, and a reliable grading system is needed to assess HE. The lack of an accurate grading system could lead to inconsistencies in the assessment of HE and poor functional prognosis and preclude comparisons of the effectiveness of treatment at different institutions.

*Address correspondence to:

Drs. Qiang Zhang and Mingfei Yang, Department of Neurosurgery, Qinghai Provincial People's Hospital, Xining 810007, Qinghai, China.

E-mail: zhangqiang691212@163.com (Zhang Q)
iloveyoucmu@163.com (Yang MF)

The purpose of this study was to devise a reliable score for prediction of HE in ICH based on several objective predictors that can be ascertained soon after onset. To that end, the current study devised such a grading system to assess the risk of HE, thus paving the way forward for standardization of HE treatment in ICH.

2. Materials and Methods

2.1. Patient selection

The Ethics Committee of Qinghai Provincial People's Hospital approved this retrospective study. Subjects were patients (> 18 years) with ICH between Jan. 2015 and Jan. 2018 who had undergone a baseline CT scan at this Hospital within 24 hours after the onset of ICH symptoms. A follow-up NCCT scan was performed within 24 hours after the baseline CT scan. At the baseline, all hematomas were located in the basal ganglia. Patients who had undergone urgent hematoma evacuation before the follow-up CT scan were excluded from this study. Also excluded were patients whose ICH was secondary to vessel malformation, head trauma, cerebral aneurysm, brain tumor, or brain infarction. The time to initial and follow-up CT scans and fundamental clinical variables were recorded for each patient. Patients with large fluctuations in blood pressure during hospitalization had their blood pressure continuously monitored while receiving standard treatment for hypertension.

All of the patients with HE were diagnosed using NCCT. The rate of HE can truly reflect basic pathological changes, so an attempt was made to devise a grading system to evaluate the risk of HE in ICH.

2.2. Imaging analysis

CT images can be analyzed in two ways: (1) imaging findings based on changes in hematoma density and shape on CT; and (2) initial hematoma volume and follow-up hematoma volume.

HE in ICH is identified based on the baseline CT scan after onset and a follow-up CT scan within 24 hours. All of the current patients were classified into 3 categories based on hematoma imaging, including changes in density and shape on CT: the island sign, blend sign, or swirl sign.

The island sign was defined as (i) ≥ 3 scattered small hematomas entirely separate from the main hematoma; (ii) or ≥ 4 small hematomas partly or all of which may connect to the main hematoma; (iii) The scattered small hematomas were round or oval and were separate from the main hematoma; (iv) The small hematomas that connected to the main hematoma were bubble-like or bud-like but not segmented (4).

The blend sign is defined as blending of a relatively

hypoattenuated area with an adjacent hyperattenuated region within the hematoma with a well-defined margin between these regions and a delta of at least 18 Hounsfield units between the 2 regions (5).

The swirl sign was defined as: hypo- or iso-density within a region of a hyperdensity that correlates with active hemorrhage on surgical evacuation (6).

HE in ICH is defined as a 33% increase in hematoma volume or > 6 mL at the time of the follow-up CT scan according to previous studies (7,8).

Two experienced observers (including an imaging physician and a neurosurgeon) who were blinded to the clinical information on patients reviewed all images to identify the 3 imaging markers. Discrepancies regarding the presence of the markers were resolved through consensus. 3D Slicer (Version 4.8.0, Harvard University, NY) was used to calculate the hematoma volume.

2.3. Statistical analysis

All statistical analyses were performed using the software package SPSS (Version 19.0; IBM Corporation, Armonk, NY). Medians and interquartile ranges (IQRs) or the mean \pm standard deviation (SD) was used to describe continuous variables, and percentage (%) was used to describe discrete variables.

The rate of HE was assessed using the t test or χ^2 test. Univariate and multivariate logistic regression analyses were used to assess predictors associated with HE. The same patient data were used to identify the accuracy of the HE prediction scores. $P < 0.05$ was considered statistically significant.

3. Results

3.1. Baseline characteristics

The characteristics of the study population are shown in Table 1. Subjects included 173 men (65.04%) and 93 women (34.96%) with a mean age of 58.98 ± 10.91 years. In the 266 patients, the hematoma was located in the basal ganglia at the baseline, including 99 (37.22%) with HE and 167 (62.78%) with non-HE. The time from onset to the initial CT scan was within 6 hours in 125 patients (46.99%) and 6-24 hours in 141 (53.01%). The overall median initial hematoma volume was 14.93 mL (IQR, 7.49-24.32 mL). Intraventricular hemorrhage (IVH) extension was noted in 83 of 266 patients (31.20%) on the baseline CT scan. Of 266 patients, 44 (16.54%) were using an anticoagulant or had an international normalized ratio (INR) > 1.5. The overall incidence of HE was 37.22% (99 of 266).

3.2. Frequency and characteristics of imaging markers

Of 1,400 potential subjects, 266 served as subjects

for analysis of HE. The island sign was noted in 48 (48.48%) of 99 patients with HE versus 13 (7.78%) of 167 patients without HE on a CT scan during patient selection. The blend sign was noted in 34 (34.34%) of 99 patients versus 29 (17.37%) of 167 patients without

HE. The swirl sign was noted in 25 (25.25%) of 99 patients versus 25 (14.97%) of 167 patients without HE. The original clinical and imaging characteristics of patients with and without CT markers are shown in Table 2.

Table 1. Baseline demographic, clinical and imaging characteristics of the population in this study

Items	Patients with HE (n = 99)	Patients without HE (n = 167)	P
Demographic characteristics			
Mean age, y (SD)	59.61 ± 10.36	58.35 ± 11.46	0.373
Sex, male, n (%)	56 (56.57)	117 (70.06)	0.725
Clinical features			
Time to baseline CT scan, Median (IQR), h	5 (4-8)	6 (5-8)	0.019
Baseline hematoma volume (IQR), mL	16.34 (8.40-29.31)	14.84 (6.98-22.63)	0.108
IVH extension at initial CT, n (%)	46 (46.46)	37 (22.16)	< 0.001
Anticoagulant use or an INR > 1.5, n (%)	27 (27.27)	17 (10.18)	< 0.001
Presence of SAH, n (%)	7 (7.07)	11 (6.59)	0.879
Presence of MLS, n (%)	18 (18.18)	26 (15.57)	0.579
Imaging markers			
Island sign, n (%)	48 (48.48)	13 (7.78)	< 0.001
Blend sign, n (%)	34 (34.34)	29 (17.37)	< 0.001
Swirl sign, n (%)	25 (25.25)	25 (14.97)	0.038

HE, hematoma expansion; CT, computed tomography; ICH, intracerebral hemorrhage; IVH, intraventricular hemorrhage; SD, standard deviation; IQR, inter-quartile range; INR, international normalized ratio.

Table 2. Comparison of baseline demographic, clinical, and imaging characteristics between patients with and without CT imaging markers

Items	IS (+), n = 61	IS (-), n = 205	P
Demographic characteristics			
Mean age, y (SD)	57.92 ± 9.50	59.09 ± 11.49	0.469
Sex, male, n (%)	46 (75.41)	127 (61.95)	0.053
Clinical features			
Baseline ICH volume (IQR), mL	24.70 (8.49-29.00)	14.51 (7.28-22.33)	< 0.01
Anticoagulant use or an INR > 1.5, n (%)	13 (29.55)	31 (70.45)	0.253
Time to baseline CT scan, median (IQR), h	5 (4-8)	6 (4-8)	0.298
IVH extension, n (%)	25 (39.68)	58 (28.29)	0.975
HE, n (%)	48 (78.69)	51 (24.88)	< 0.001
Items	BS (+), n = 63	BS (-), n = 203	P
Demographic characteristics			
Mean age, y (SD)	57.30 ± 11.54	59.29 ± 10.89	0.213
Sex, male, n (%)	45 (71.43)	128 (63.05)	0.223
Clinical features			
Baseline ICH volume (IQR), mL	20.78 (13.95-32.96)	12.21 (6.00-22.03)	< 0.01
Anticoagulant use or an INR > 1.5, n (%)	13 (29.55)	31 (70.45)	0.317
Time to baseline CT scan, median (IQR), h	6 (4-9)	5 (4-8)	0.212
IVH extension, n (%)	20 (31.75)	63 (31.03)	0.915
HE, n (%)	34 (53.97)	65 (32.02)	0.002
Items	SS (+), n = 50	SS (-), n = 216	P
Demographic characteristics			
Mean age, y (SD)	59.40 ± 10.20	58.69 ± 11.24	0.681
Sex, male, n (%)	31 (62.00)	142 (65.74)	0.617
Clinical features			
Baseline ICH volume (IQR), mL	19.89 (11.82-28.28)	14.23 (7.12-22.79)	0.017
Anticoagulant use or an INR > 1.5, n (%)	9 (20.45)	35 (79.55)	0.758
Time to baseline CT scan, median (IQR), h	5 (3-7)	6 (4-8)	0.024
IVH extension, n (%)	15 (30.00)	68 (31.48)	0.839
HE, n (%)	25 (50.00)	74 (34.26)	0.038

IS, island sign; BS, blend sign; SS, swirl sign.

The presence of imaging markers was associated with a larger initial hematoma volume, the island sign (24.70 mL: 14.51 mL; $P < 0.01$), the blend sign (20.78 mL: 12.21 mL; $P < 0.01$), and the swirl sign (19.89 mL: 14.23 mL; $P = 0.017$). Anticoagulant use or an INR > 1.5 , the time from onset to the baseline CT scan on admission, and IVH extension did not differ significantly in patients with or without imaging markers.

The 2 observers identified imaging markers with a high level of inter-observer agreement ($\kappa = 0.90$). The island sign predicted HE with a sensitivity of 48.48%, a specificity of 92.22%, a positive predictive value of 78.69%, and a negative predictive value of 75.12%. The blend sign predicted HE with a sensitivity of 34.34%, a specificity of 82.63%, a positive predictive value of 53.97%, and a negative predictive value of 67.98%. The swirl sign predicted HE with a sensitivity of 25.25%, a specificity of 85.03%, a positive predictive value of 50.00%, and a negative predictive value of 65.74%. All of the imaging markers studied had satisfactory ability to predict HE. These markers differed significantly in the sensitivity, specificity, and NPV with which they predicted HE (all $P < 0.05$). The island sign had better ability to predict HE. Relevant data are shown in Table 3.

Univariate and multivariate logistic regression analysis were performed to assess the association

between various clinical and imaging parameters and HE, as shown in Table 4. Univariate analysis indicated that the time to the baseline CT scan (odds ratio (OR), 1.064; 95% confidence interval (CI), 1.003-1.127; $P = 0.038$), initial ICH volume (OR, 1.021; 95% CI, 1.003-1.038; $P = 0.018$), IVH extension (OR, 3.049; 95% CI, 1.781-5.222; $P < 0.001$), anticoagulant use or an INR > 1.5 (OR, 3.309; 95% CI, 1.695-6.458; $P < 0.001$), and the presence of imaging markers (island sign, blend sign, or swirl sign) on a CT scan upon admission (all $P < 0.005$) were associated with HE. Multivariate logistic regression analysis indicated that the time to the baseline CT scan (OR, 0.986; 95% CI, 0.926-1.014; $P < 0.001$), the baseline ICH volume (OR, 0.974; 95% CI, 0.949-0.999; $P = 0.042$), anticoagulant use or an INR > 1.5 (OR, 3.362; 95% CI, 1.415-7.988; $P = 0.006$), and the presence of imaging markers on the baseline CT scan ($P < 0.001$ for all) independently predicted HE. The results of logistic regression analysis for HE are shown in Table 4.

3.3. Devising and validation of the grading system

A grading system was devised using the parameters from multivariable regression in Table 4, including imaging markers (island sign, blend sign, or swirl sign), the time from onset to the initial CT scan, anticoagulant

Table 3. Sensitivity, specificity, PPV, and NPV of CT imaging markers

Items	Island sign (95% CI)	Blend sign (95% CI)	Swirl sign (95% CI)	P
Sensitivity	48.48 (42.48-54.49)	34.34 (28.64-40.05)	25.25 (20.03-30.47)	0.028
Specificity	92.22 (89.00-95.44)	82.63 (78.08-87.19)	85.03 (80.74-89.32)	0.003
PPV	78.69 (73.77-83.61)	53.97 (47.98-59.96)	50.00 (43.99-56.01)	0.094
NPV	75.12 (69.93-80.32)	67.98 (62.37-73.59)	65.74 (60.04-71.44)	0.003

PPV, positive predictive value; NPV, negative predictive value.

Table 4. Univariate and multivariate analysis of predictors for HE

Items	OR	95% CI	P
Univariate analysis			
Age	1.010	0.998-1.033	0.372
Time to baseline CT scan	1.064	1.003-1.127	0.038
Baseline ICH volume	1.021	1.003-1.038	0.018
IVH extension	3.049	1.781-5.222	< 0.001
Anticoagulant use or an INR > 1.5	3.309	1.695-6.458	< 0.001
Presence of imaging markers			
Island sign	11.149	5.593-22.224	< 0.001
Blend sign	2.489	1.399-4.430	0.003
Swirl sign	1.919	1.031-3.573	0.042
Multivariate analysis			
Time to baseline CT scan	0.986	0.926-1.014	< 0.001
Baseline ICH volume	0.974	0.949-0.999	0.042
Anticoagulant use or an INR > 1.5	3.362	1.415-7.988	0.006
IVH extension	3.225	1.501-6.929	0.003
Presence of imaging markers			
Island sign	39.503	15.022-103.882	< 0.001
Blend sign	15.300	5.933-39.456	< 0.001
Swirl sign	9.798	3.785-25.368	< 0.001

CI, confidence interval; OR, odds ratio.

Table 5. Summary of the HE Prediction Grading System

Component	Points
Hours from onset to CT (h)	1
≤ 6	0
6-24	
Baseline ICH volume (mL)	1
≥ 30	0
< 30	
Island sign	1
Present	0
Absent	
Blend sign	1
Present	0
Absent	
Swirl sign	1
Present	0
Absent	
Anticoagulant use or an INR > 1.5	1
Present	0
Absent	
IVH extension	1
Present	0
Absent	

Total points:0-7.

use or an INR > 1.5, and IVH extension. The grading system is shown in Table 5. These factors were used to create a statistical model to predict HE.

Based on the results of the regression model, CT imaging markers (island sign, blend sign, or swirl sign), anticoagulant use or an INR > 1.5, and time from onset to the initial CT scan were the strongest predictors of HE. The results of these regression models have laid the foundation for scoring the prediction of HE in ICH.

Therefore, the score to predict HE in ICH consists of 5 components: changes in density and shape on CT (imaging markers), the time from onset to the initial CT scan, anticoagulant use or an INR > 1.5, baseline ICH volume, and IVH extension. The time from onset to the initial CT scan was further subdivided (≤ 6 h and 6-24 h), as was the baseline ICH volume (≥ 30 mL and < 30 mL). A baseline ICH volume ≥ 30 mL had a sensitivity of 32.32% and a specificity of 80.84%. Each variable was given a specific cutoff based on its relevance to HE. A baseline CT scan performed within 6 hours of onset is critical because it represents the critical nature of HE. The grading system score ranges from 0 to 7 points.

The grading system was used to score patients in order to determine its ability to predict HE, and those results are shown in Table 6. The incidence of HE increased as the score increased. A higher score on the grading system indicated a greater probability of HE ($P < 0.001$). The probability of HE was 3.45% for a score of 0 (1/29), 18.31% (13/71) for a score of 1, 32.53% (27/83) for a score of 2, 61.22% (30/49) for a score of 3, 79.17% (19/24) for a score of 4, and 85.71% (6/7) for a score of 5. The probability of HE was 100% (3/3) with a score ≥ 6. No patients had a score of 7. The grading

Table 6. Risk of HE according to the Prediction Grading System

Total Score	Estimate of HE risk, % (n)
0	3.45 (1/29)
1	18.31 (13/71)
2	32.53 (27/83)
3	61.22 (30/49)
4	79.17 (19/24)
5	85.71 (6/7)
≥ 6	100 (3/3)

system's performance in predicting HE prediction is shown in Table 6.

3.4. Representative case

A 56-year-old man with hypertension and diabetes mellitus was first admitted to this Hospital for weakness in his right extremities secondary to a left basal ganglia hematoma within 4 hours of the onset of symptoms. A baseline CT scan revealed a combined left basal ganglia hematoma. Two experienced observers (including an imaging physician and a neurosurgeon) who were blinded to the clinical information on the patient reviewed the CT images, and both noted the blend sign and the swirl sign (Figure 1, A). 3D Slicer was used to calculate the baseline ICH volume, which was about 73.29 mL. The patient had been prescribed warfarin for atrial fibrillation. His blood pressure upon admission was 232/103 mmHg, and his Glasgow Coma Scale (GCS) score was 5. His INR upon admission was 1.8. According to the grading system, the patient's score was 5 points (hours from onset to CT = 4; baseline ICH volume = 73.29 mL; presence of the blend sign and swirl sign; and anticoagulant use and INR = 1.8). Accordingly, the patient was deemed to have a high risk of HE. A follow-up NCCT scan was performed 18 hours after onset. 3D-Slicer was again used to measure hematoma volume, which was about 84.50 mL (Figure 1, B).

4. Discussion

Although HE is a common phenomenon in ICH, there is no widely accepted grading system that can be used to predict HE in patients with ICH and to guide clinical treatment and research. A lack of HE prediction scores has presumably lead to large differences in the clinical data collected in clinical studies of ICH and inconsistent treatment strategies. A new grading system has been devised to predict the risk of HE in patients with ICH by analyzing clinical data on patients with ICH at this Hospital. The significance for each component of this grading system is worth discussing. The predictive grading system has five components: the time from onset to the initial CT scan, baseline ICH volume, the presence of imaging markers (island sign, blend sign, or

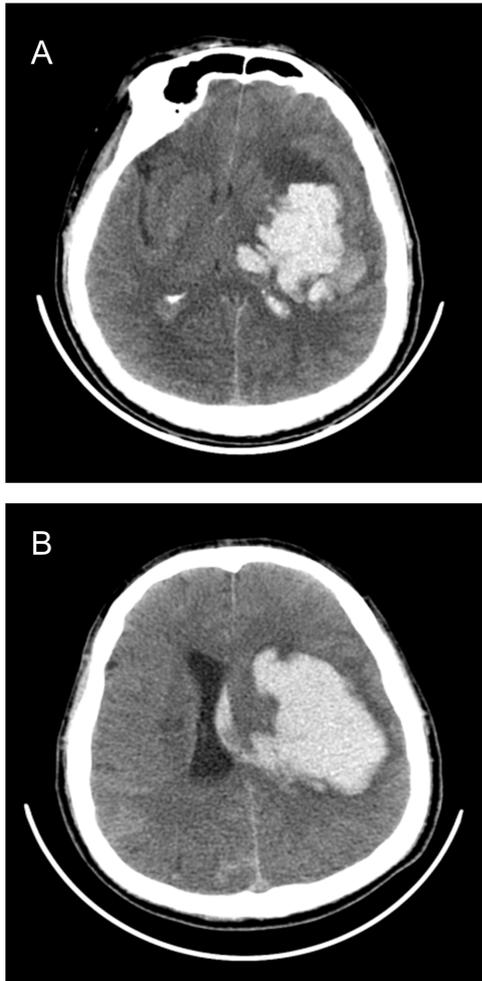


Figure 1. A CT scan showing a combined left basal ganglia hematoma (Figure 1, A) when the patient developed motor weakness in his right extremities during initial admission. A CT scan on second admission revealed a larger left basal ganglia hematoma (Figure 1, B).

swirl sign), anticoagulant use or an INR > 1.5, and IVH extension. This system provides an easily performed method of reasonably predicting HE.

Some of the components of the grading system have been previously cited as independent predictors of HE in ICH. The time from onset to the initial CT scan has been identified as a consistent predictor of HE in various studies (9,10,12-15). HE occurs soon after the onset of ICH, usually within 3-6 h, in approximately 1/4 to 1/3 of patients (10). In the current study, the time from onset to the initial CT scan was within 6 hours in 125 patients (46.99%) and 6-24 hours in 141 (53.01%). In this study, the time to the baseline CT scan was significantly associated with HE ($P < 0.001$).

In patients with ICH, the presence of NCCT imaging markers in the form of the island sign, the blend sign and/or the swirl sign indicates a greater possibility of HE (4-6). The presence of imaging markers has been identified as an independent predictor of HE, and these markers have a reliable accuracy at predicting HE (4-6,23). Therefore, these markers could

be used to predict HE in ICH. In the current study, the sensitivity, specificity, and NPV of these markers differed significantly ($P < 0.05$ for all). The island sign had better ability to predict HE.

A causal link between anticoagulant use and HE seems logical, but relevant studies have yielded inconsistent results (7,10,13,15-18). In the current study, anticoagulant use or an INR > 1.5 was significantly associated with HE ($P < 0.001$). IVH extension and anticoagulant use are also associated with HE according to several previous studies (11-15). Moreover, these two variables had been identified as significant predictors of HE in previous studies (4,19,20). In the current study, these two risk factors were highly associated with HE (both $P < 0.001$). Therefore, they were included in the grading system developed here. A baseline ICH volume larger than 30 mL has been found to be a risk factor for HE according to a couple of studies (9,11). The current study yielded similar results (OR, 0.974; $P = 0.042$). Therefore, baseline ICH volume was included as a predictive parameter with a sensitivity of 32.32% and a specificity of 80.84% (ICH volume ≥ 30 mL).

Recently, a study by Li *et al.* (4) found that a lower GCS score is one of critical predictors that can influence HE ($P < 0.001$). Because of differences in anatomy and blood supply between basal ganglia and non-basal ganglia, the current study examined ICH in the basal ganglia in order to reduce bias. Moreover, the GCS was not included as a parameter since all of the patients with ICH had a hematoma located in the basal ganglia. In most people, one hemisphere of the brain is dominant (left or right). When the level of consciousness or ability to speak is assessed in patients with ICH, their GCS will inevitably be inaccurate, and this might influence the results of a grading system.

Brouwers *et al.* (9) developed a 9-point score to predict HE based on four parameters: the presence of the spot sign, warfarin use, the time to the initial CT (> 6 h or ≤ 6 h), and baseline ICH volume (< 30 mL, 30-60 mL or > 60 mL). In their study, a higher score resulted in better ability to predict HE. In 2015, Wang *et al.* (21) refined the 9-point score by including baseline ICH volume (≤ 10 mL, 10-20 mL or > 20 mL) and the time to the initial CT (≤ 1 h, 1-2 h, 2-3 h, 3-4 h, 4-5 h and > 5 h). Based on the 9-point score, Wang *et al.* (19) proposed a new 24-point score (BRAIN) and they added two novel parameters: IVH extension and recurrent ICH. The scores from that system are similar to those in previous studies. The higher the score, the greater the probability of HE. Although the accuracy of these two system seems comparable to that of the grading system developed here, this new grading system has several additional advantages. Because of the universality of NCCT, the parameters needed to tally a score with this grading system are readily determined in almost any medical facility. More importantly, imaging makers (island sign, blend sign and/or swirl sign) can be readily

identified in clinical settings. That said, the study by Brouwers *et al.* (9) included a larger number of patients with ICH >30 mL, suggesting that the two grading systems might be complementary at predicting HE depending on clinical characteristics.

The predictive grading system devised here can provide a reasonable estimation of the risk for HE in patients with ICH, and it could have useful clinical applications. The grading system indicated that a higher score meant a greater probability of HE ($P < 0.001$). The probability of HE was 3.45% (1/29) for a score of 0, 18.31% (13/71) for a score of 1, 32.53% (27/83) for a score of 2, 61.22% (30/49) for a score of 3, 79.17% (19/24) for a score of 4, and 85.71% (6/7) for a score of 5. The probability of HE was 100% (3/3) for a score \geq 6. This score may help neurosurgeons to decide which patients with ICH will require closer monitoring or surgery. It may also aid in identifying patients who would benefit most from interventions targeting HE in clinical trials (22).

The current study had several limitations. First, this study was retrospective in nature and conducted at a single institution with a relatively small sample size, so the current results need to be verified at other medical facilities. Second, a relationship between HE and neurological deterioration or long-termed clinic functional outcomes was not evident, so a scoring threshold could not be defined to predict clinically significant HE. Future prospective studies need to be conducted to define appropriate thresholds to stratify patients with HE a low versus a high risk of clinical deterioration to guide clinical decision-making. Last, the grading system developed here has not been validated in separate internal or external cohorts. This limits the generalizability of the current results to other patients with ICH. Therefore, the grading system developed here should not be used to make final treatment recommendations or to obtain definitive information on the progression of ICH or prognosis at this stage in time. Future plans are to validate the grading system developed here in an external cohort of patients with ICH.

In conclusion, a quick and easy-to-use grading system was developed and internally validated to predict the risk for HE in a cohort of patients with ICH. This predictive grading system could have major clinical applications.

Acknowledgements

This study was funded by grants from the Qinghai Province Health and Family Planning Commission (2016WJZDX13).

References

- Brouwers HB, Greenberg SM. Hematoma expansion following acute intracerebral hemorrhage. *Cerebrovasc Dis.* 2013; 35:195-201.
- Qureshi AI, Mendelow AD, Hanley DF. Intracerebral hemorrhage. *Lancet.* 2009; 373:1632-1644.
- Flaherty ML, Haverbusch M, Sekar P, Kissela B, Kleindorfer D, Moomaw CJ, Sauerbeck L, Schneider A, Broderick JP, Woo D. Long-term mortality after intracerebral hemorrhage. *Neurology.* 2006; 66:1182-1186.
- Li Q, Liu QJ, Yang WS, Wang XC, Zhao LB, Xiong X, Li R, Cao D, Zhu D, Wei X, Xie P. Island Sign: An imaging predictor for early hematoma expansion and poor outcome in patients with intracerebral hemorrhage. *Stroke.* 2017; 48:3019-3025.
- Li Q, Zhang G, Huang YJ, Dong MX, Lv FJ, Wei X, Chen JJ, Zhang LJ, Qin XY, Xie P. Blend sign on computed tomography: Novel and reliable predictor for early hematoma growth in patients with intracerebral hemorrhage. *Stroke.* 2015; 46:2119-2123.
- Selariu E, Zia E, Brizzi M, Abul-Kasim K. Swirl sign in intracerebral hemorrhage: Definition, prevalence, reliability and prognostic value. *BMC Neurol.* 2012; 12:109.
- Selariu E, Zia E, Brizzi M, Abul-Kasim K. PREDICT/Sunnybrook ICH CTA Study Group. Prediction of hematoma growth and outcome in patients with intracerebral hemorrhage using the CT-angiography spot sign(PREDICT): A prospective observational study. *Lancet Neurol.* 2012; 11:307-314.
- Dowlatshahi D, Demchuk AM, Flaherty ML, Ali M, Lyden PL, Smith EE; VISTA Collaboration. Defining hematoma expansion in intracerebral hemorrhage: Relationship with patient outcomes. *Neurology.* 2011; 76:1238-1244.
- Brouwers HB, Chang Y, Falcone GJ, *et al.* Predicting hematoma expansion after primary intracerebral hemorrhage. *JAMA Neurol.* 2014; 71:158-164.
- Toyoda K, Okada Y, Minematsu K, Kamouchi M, Fujimoto S, Ibayashi S, Inoue T. Antiplatelet therapy contributes to acute deterioration of intracerebral hemorrhage. *Neurology.* 2005; 65:1000-1004.
- Romero JM, Brouwers HB, Lu J, Delgado Almandoz JE, Kelly H, Heit J, Goldstein J, Rosand J, Gonzalez RG. Prospective validation of the computed tomographic angiography spot sign score for intracerebral hemorrhage. *Stroke.* 2013; 44:3097-3102.
- Dowlatshahi D, Smith EE, Flaherty ML, Ali M, Lyden P, Demchuk AM; VISTA Collaborators. Small intracerebral haemorrhages are associated with less haematoma expansion and better outcomes. *Int J Stroke.* 2011; 6:201-206.
- Cucchiara B, Messe S, Sansing L, Kasner S, Lyden P; CHANT Investigators. Hematoma growth in oral anticoagulant related intracerebral hemorrhage. *Stroke.* 2008; 39:2993-2996.
- Broderick JP, Dinger MN, Hill MD, Brun NC, Mayer SA, Steiner T, Skolnick BE, Davis SM; Recombinant activated Factor VII Intracerebral Hemorrhage Trial Investigators. Determinants of intracerebral hemorrhage growth: An exploratory analysis. *Stroke.* 2007; 38:1072-1075.
- Rodriguez-Luna D, Piñeiro S, Rubiera M, Ribo M, Coscojuela P, Pagola J, Flores A, Muchada M, Ibarra B, Meler P, Sanjuan E, Hernandez-Guillamon M, Alvarez-Sabin J, Montaner J, Molina CA. Impact of blood pressure changes and course on hematoma growth in acute intracerebral hemorrhage. *Eur J Neurol.* 2013;

- 20:1277-1283.
16. Rodriguez-Luna D, Rubiera M, Ribo M, Coscojuela P, Pagola J, Piñeiro S, Ibarra B, Meler P, Maisterra O, Romero F, Alvarez-Sabin J, Molina CA. Serum low-density lipoprotein cholesterol level predicts hematoma growth and clinical outcome after acute intracerebral hemorrhage. *Stroke*. 2011; 42:2447-2452.
 17. Kuramatsu JB, Mauer C, Kipphuth IC, Lücking H, Kloska SP, Köhrmann M, Huttner HB. Reported antiplatelet use influences long-term outcome independently in deep intracerebral hemorrhage. *Neurosurgery*. 2012; 70:342-350.
 18. Yildiz OK, Arsava EM, Akpınar E, Topcuoglu MA. Previous antiplatelet use is associated with hematoma expansion in patients with spontaneous intracerebral hemorrhage. *J Stroke Cerebrovasc Dis*. 2012; 21:760-766.
 19. Huynh TJ, Aviv RI, Dowlatshahi D, *et al*. Validation of the 9-Point and 24-Point hematoma expansion prediction scores and derivation of the PREDICT A/B scores. *Stroke*. 2015; 46:3105-3110.
 20. Hokari M, Shimbo D, Asaoka K, Uchida K, Itamoto K. Impact of antiplatelets and anticoagulants on the prognosis of intracerebral hemorrhage. *J Stroke Cerebrovasc Dis*. 2018; 27:53-60.
 21. Wang X, Arima H, Al-Shahi Salman R, Woodward M, Heeley E, Stapf C, Lavados PM, Robinson T, Huang Y, Wang J, Delcourt C, Anderson CS; INTERACT Investigators. Clinical prediction algorithm (BRAIN) to determine risk of hematoma growth in acute intracerebral hemorrhage. *Stroke*. 2015; 46:376-381.
 22. Yao X, Xu Y, Siwila-Sackman E, Wu B, Selim M. The HEP Score: A Nomogram-derived hematoma expansion prediction scale. *Neurocrit Care*. 2015; 23:179-187.
 23. Li R-L, Yang M-F. A comparative study of the blend sign and the black hole sign on CT as a predictor of hematoma expansion in spontaneous intracerebral hemorrhage. *BioSci Trends*. 2017; 11:682-687.
- (Received January 3, 2018; Revised April 12, 2018; Accepted April 25, 2018)

A simple and economical method of gas chromatography-mass spectrometry to determine the presence of 6 pesticides in human plasma and its clinical application in patients with acute poisoning

Guiyan Yuan, Rui Zhang, Xuwang Chen, Benjie Wang, Ruichen Guo*

Institute of Clinical Pharmacology, Qilu Hospital of Shandong University, Jinan Shandong, China.

Summary

An economical, rapid, and sensitive method of gas chromatography-mass spectrometry (GC-MS) was developed and validated to determine the presence of six pesticides (dichlorvos, acetochlor, atrazine, chlorpyrifos, α -endosulfan, and β -endosulfan) in human plasma. The pesticides were extracted with acetonitrile and concentrated using anhydrous sodium sulfate. Then, the target compounds were analyzed and quantified with GC-MS using borneol as an internal standard. Separation was performed on a HP-5MS capillary column (30 m \times 0.25 mm \times 0.25 μ m) with temperature programming. Detection was accomplished under electro-spray ionization (ESI) in selected ion monitoring (SIM) mode. Under optimized conditions, satisfactory linear ranges of 0.05-10 μ g/mL were obtained for all of the analyzed pesticides. The linear correlation coefficients were greater than 0.99. The average recovery was between 86.8 and 106.5%. The inter- and intra-day precision ranged from 1.7-14.5% and 4.2-13.8%, respectively. Dichlorvos was unstable in plasma both at room temperature and when frozen. The other five pesticides were stable after storage at -20°C for 17 days and two freeze-thaw cycles. Thirty-five plasma samples from 15 patients with acute self-poisoning were analyzed using this method. Dichlorvos was found in 13 plasma samples with a mean concentration of 0.289 μ g/mL, and atrazine was found in 6 with a mean concentration of 0.261 μ g/mL. Acetochlor was found in one plasma sample (0.153 μ g/mL). This method is simple, reliable and cost-effective. It takes little time and does not waste solvents, and it can be used to routinely detect six pesticides in patients with acute poisoning.

Keywords: Gas chromatography-mass spectrometry, pesticides, determination, plasma

1. Introduction

Numerous pesticides have been widely used in agriculture, industry, and medicine. However, the long-term and large-scale use of these chemicals inevitably contaminates the environment, thereby posing a serious threat to human health (1,2). Pesticides are leading causes of morbidity and mortality following intentional self-poisoning or in cases of occupational or environmental exposure (3,4). Self-poisoning with pesticides is a major problem (5). Thus, sensitive and

efficient methods to detect pesticides in human blood or urine need to be developed.

Rapid identification and quantification of causal pesticides would provide useful information to clinicians so that they can make appropriate treatment decisions (6). In cases of acute poisoning, determining the presence of a pesticide in human blood can provide information to identify the type of pesticide, the absorbed dose, and the degree of exposure to target tissues prior to elimination from the body (7,8). The volatility, thermal stability, and low polarity of pesticides render them suitable for gas chromatographic analysis, and particularly for the determination of their presence in biological matrices (9). Gas chromatography coupled with mass spectrometry (GC-MS or GC-MS/MS), which can avoid most matrix interference, has been used when a highly selective

*Address correspondence to:

Dr. Ruichen Guo, Institute of Clinical Pharmacology, Qilu Hospital of Shandong University, No. 107 Wenhua Xi Road, Jinan 250012, Shandong Province, China.
E-mail: grc7636@126.com

detection is required (10,11).

A few methods of gas chromatography with specific detectors including a flame ionization detector (FID) (12,13), a nitrogen-phosphorus detector (NPD) (14,15), a flame photometric detector (FPD) (16,17), and mass spectrometer detectors (18-20) have been used to determine the presence of pesticides in human blood, plasma, or serum. However, these methods were mainly developed to determine the presence of organophosphate pesticides and not other types of pesticides, and the extraction procedure used was mainly solid phase extraction (SPE), which is complex, takes time, and is expensive. A simple and rapid method of sample extraction needs to be developed to determine the presence of pesticide in order to provide useful information to clinicians for accurate diagnosis and treatment.

The aim of this study was to develop, optimize, and validate a simple and rapid method for simultaneous determination of the presence of six pesticides including dichlorvos, acetochlor, atrazine, chlorpyrifos, α -endosulfan, and β -endosulfan in human plasma. These pesticides different from those detected using the methods mentioned above, and they were selected based on the frequency of cases of poisoning, clinical requirements, and their physical and chemical properties. Hence, a new and sensitive method for simultaneous determination of the presence of these pesticides in human plasma was developed using a combination of rapid protein precipitation (PPT) extraction and GC-MS. The method is simple, rapid, economical, and successfully detects the selected pesticides in patients with acute self-poisoning.

2. Materials and Methods

2.1. Reagents and materials

Dichlorvos (99.2%, Lot No. 20161121), acetochlor (99.2%, Lot No. 20161125), atrazine (99.4%, Lot No. 20161207), and chlorpyrifos (98.0%, Lot No. 20160810) were purchased from Shanghai Pesticide Research Institute Co., Ltd. (Shanghai, China). α -endosulfan (99.0%, Lot No. 41020) and β -endosulfan (99.0%, Lot No. 41013) were purchased from Dr. Ehrenstorfer GmbH (Augsburg, Germany). Borneol (internal standard, Lot No. 110881-200706) was provided by the National Institutes for Food and Drug Control (Beijing, China). Chromatographic-grade acetonitrile and hexanes were purchased from J.T. Baker (Phillipsburg, New Jersey, USA). Anhydrous sodium sulfate (99%, Lot C1410040) was purchased from Aladdin Industrial Corporation (Shanghai, China). Blank human plasma was collected from healthy donors and stored immediately at -20°C until analysis. This study was approved by the ethics committee of Qilu Hospital of Shandong University. Signed informed

consent was obtained from each donor.

2.2. Instrumentation and analytical conditions

The GC apparatus used in this study was an Agilent 7890A equipped with an Agilent 7683 autosampler, and the split/splitless injector was operated in splitless mode. The apparatus was coupled to an Agilent 5975C mass spectrometer (Agilent Technologies, Inc, Santa Clara, CA, USA). The analytical column was an HP-5MS Agilent column (30 m \times 0.25 mm \times 0.25 μm , Santa Clara, CA, USA). The chromatograph was programmed for an initial temperature of 80°C for 4 min. The temperature was increased to 200°C at a rate of $70^{\circ}\text{C}/\text{min}$ and then increased to 250°C at a rate of $5^{\circ}\text{C}/\text{min}$, held for 2 min, and then increased to 300°C at a rate of $50^{\circ}\text{C}/\text{min}$, held for 5 min. The total run time was 23.7 min. Helium was used as the carrier gas (flow rate: 2 mL/min). The mass spectrometer was operated in the electron impact (70 eV) and selected ion monitoring (SIM) mode both for qualitative and for quantitative analysis, with a solvent delay of 3.75 min. For each analyte, the most abundant and characteristic mass fragment ion was chosen for quantitation along with two others for confirmation, as shown in Table 1. The compounds were subsequently identified based on their relative retention times and the ratio of their respective confirmation ions to their quantitation ion.

2.3. Preparation of standard solutions, calibration standards, and quality control samples

A stock solution for each pesticide was prepared at a concentration of 2 mg/mL in acetonitrile. Mixed working solutions of the six pesticides were prepared at concentrations of 1, 2, 4, 10, 20, 40, 100, and 200 $\mu\text{g}/\text{mL}$ in acetonitrile. Calibration standards and quality control (QC) samples were prepared by appropriate dilution of the above working solutions in blank plasma to obtain concentrations of 0.05, 0.1, 0.2, 0.5, 1, 2, 5, and 10 $\mu\text{g}/\text{mL}$, and 0.15, 0.8, and 8 $\mu\text{g}/\text{mL}$. An IS solution of borneol was prepared at 1 mg/mL in acetonitrile and further diluted with acetonitrile to reach a concentration of 3 $\mu\text{g}/\text{mL}$. All stock and working solutions in acetonitrile were stored at 4°C .

Table 1. GC parameters used for the selected pesticides

Pesticide	Retention time (min)	Quantitation ion (m/z)	Confirmation ions (m/z)
Borneol	6.00	95.1	110.1, 138.2
Dichlorvos	6.49	109.0	185.0, 79.0
Atrazine	9.12	314.0	196.9, 257.9
Acetochlor	10.17	146.1	162.1, 223.1
Chlorpyrifos	11.11	200.0	215.1, 173.0
α -Endosulfan	12.75	194.9	240.9, 264.9
β -Endosulfan	14.16	194.9	240.9, 264.9

2.4. Sample extraction

A 200- μ L aliquot of plasma was transferred to a labeled conical plastic Eppendorf tube. Ten μ L of the IS (3 μ g/mL) solution and 200 μ L of acetonitrile were added, vortexed for 2 min, and then centrifuged at 10,800 rpm for 5 min. The supernatant portion was transferred into another labeled conical plastic Eppendorf tube containing 250 mg of anhydrous sodium sulfate, vortexed for 2 min, and then centrifuged at 10,800 rpm for 5 min. The supernatant portion was transferred into a glass vial and 1 μ L was introduced into the GC-MS system.

2.5. Method validation

The method was validated in terms of selectivity, extraction recovery, calibration curves, the lower limit of quantitation (LLOQ), precision and accuracy, carryover, and stability according to the US Food and Drug Administration and Chinese State Food and Drug Administration guidelines for the validation of bioanalytical methods.

Selectivity was studied by analyzing six blank plasma samples and plasma samples spiked with six pesticides to identify the potential interference of endogenous substances at the retention time of each pesticide and the IS under the GC-MS conditions described.

Extraction recovery was investigated at two concentrations (0.15 and 8 μ g/mL) in triplicate for each pesticide. Recovery was calculated by dividing the peak area for the pre-spiked sample by the peak area for the post-spiked sample and multiplying by 100%.

A calibration curve was generated for each pesticide using eight non-zero concentrations, typically described by the equation $y = ax + b$, where y corresponds to the peak area ratio and x to the ratio of the concentration of each pesticide to that of the IS. The calibration curves were obtained using quadratic least-squares regression with the reciprocal of the concentration squared ($1/x^2$) as the weighting factor. The LLOQ of each pesticide was defined as the lowest concentration on the calibration curve with a precision of $< 20\%$, expressed as the relative standard deviation (RSD) and an accuracy between 80 and 120%.

The intra-day precision was assessed at 0.05, 0.15, 0.8, and 8 μ g/mL through extraction and analysis of five replicates for each concentration on the same day. The inter-day precision was studied at the same concentrations on three consecutive days. The accuracy was determined by comparing the mean measured concentration to its theoretical value, and the result was expressed as the mean relative error (RE). Precision was expressed as the RSD. The RE and RSD should be less than $\pm 20\%$ for a concentration of 0.05 μ g/mL and $\pm 15\%$ for concentrations of 0.15, 0.8, and 8 μ g/mL.

Stability was studied using two concentrations

(0.15 and 8 μ g/mL) in triplicate stored or processed under different conditions, *i.e.* storage at -20°C for 17 days, freezing (-20°C) and thawing (room temperature) for two cycles, leaving plasma samples to stand on the bench top for 2 h, and leaving the post extraction samples in the auto sampler for 24 h at room temperature. The accuracy expressed as RE% should be less than $\pm 15\%$.

2.6. Application

This method was used to determine the concentrations of pesticides in the plasma of patients with suspected poisoning seen in the Emergency Department of Qilu Hospital of Shandong University between May and September 2017. Three mL of venous blood was collected into vacutainer tubes containing disodium ethylenediamine tetraacetic acid (EDTA-Na) as the anticoagulant. The tubes were centrifuged at 5,000 rpm for 5 min. The plasma was removed and placed in an Eppendorf tube and stored at -20°C until analysis. This study was carried out in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). All protocols were reviewed and approved by the scientific research ethics committee of Qilu Hospital of Shandong University. Signed informed consent forms were obtained from all patients who participated in this study or their caregivers.

3. Results and Discussion

3.1. Optimization of GC-MS conditions

Gas chromatography conditions were optimized by changing the initial temperature and heating rate. When the initial temperature was set at 50°C , dichlorvos and atrazine produced markedly tailed peaks. This situation was improved by increasing the initial temperature to 80°C . Heating had to progress at a reasonable rate, or else the separation and response of the pesticides would be seriously affected. Caffeine and borneol were tried as ISes, but there was endogenous substance interference at the retention time of caffeine. Ultimately, borneol was used as the IS because of the appropriate retention time and lack of endogenous interference.

3.2. Development of a procedure for sample extraction

The choice of a procedure for sample extraction started with liquid-liquid extraction with ethyl acetate as the extraction solvent. After centrifugation, the supernatant was transferred into a tube and dried in a water bath under a gentle N_2 flow at 40°C . Then the residue was re-dissolved using acetonitrile and analyzed using GC-MS. However, no or a very small amount of dichlorvos was found at concentration of 0.1 or 0.2 μ g/mL depending on the drying time. When protein

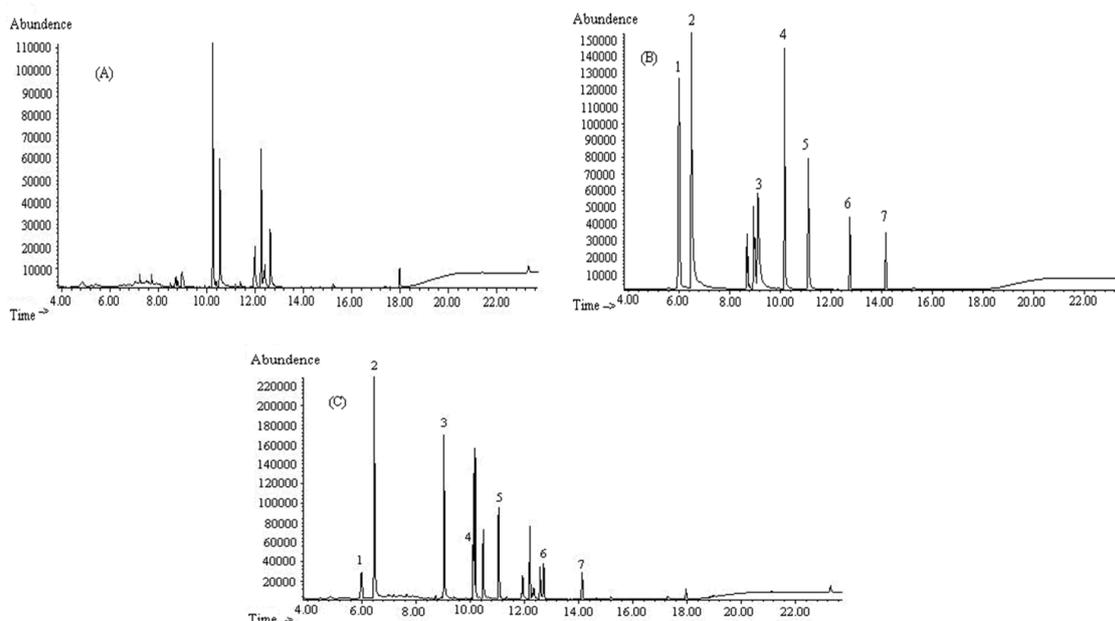


Figure 1. The total ion current chromatograms of blank plasma (A), six standard pesticides and the IS (B), and blank plasma spiked with six pesticides and the IS (C). (1-Borneol; 2-Dichlorvos; 3-Atrazine; 4-Aceto chlor; 5-Chlorpyrifos; 6- α -Endosulfan; 7- β -Endosulfan).

precipitation with acetonitrile as precipitant was used, dichlorvos was detected, and the peak areas were concentration-dependent. This indicates that dichlorvos readily volatilizes when dried with N_2 . Thus, protein precipitation was selected as the method for sample preparation. In order to reduce the LLOQ, anhydrous sodium sulfate was used for dehydration. With this method of sample preparation, the LLOQs for the six pesticides reached 0.05 $\mu\text{g/mL}$, and the extraction recovery ranged from 86.8%-106.5%. This method was simpler, faster, and more cost-effective than solid phase extraction, which is often used in literature related to the trace analysis of pesticides.

3.3. Method validation

All spiked and blank samples were free of co-eluting peaks at the retention times of the pesticides and the IS. The total ion current chromatograms of blank plasma, six standard pesticides and the IS, and blank plasma spiked with six pesticides and the IS are shown in Figure 1. There were no interfering peaks for all selected ions from the sample matrix (absence of any interfering peaks). This indicated that the method was selective to the six pesticides.

The extraction recovery of the six pesticides and the IS is shown in Table 2. In general, an average recovery greater than 86.8% was obtained, and the recovery was often close to 100%. This indicates that the recovery was good and reproducible, so it was thus deemed fit for purpose.

Calibration curves ranging from 0.05 to 10 $\mu\text{g/mL}$ were obtained for all of the pesticides. Linearity was noted with a correlation coefficient (r^2) greater than 0.99.

Table 2. The extraction recovery (%) for the six pesticides

Pesticide	0.15 $\mu\text{g/mL}$	8 $\mu\text{g/mL}$
Dichlorvos	96.3 \pm 2.9	97.9 \pm 4.4
Atrazine	106.5 \pm 13.9	94.1 \pm 6.4
Aceto chlor	104.2 \pm 11.5	96.0 \pm 6.1
Chlorpyrifos	103.9 \pm 10.0	92.6 \pm 5.9
α -Endosulfan	102.6 \pm 12.5	86.8 \pm 5.7
β -Endosulfan	102.3 \pm 4.3	88.7 \pm 6.7

The regression equation, correlation coefficient, and RSDs for LLOQ are shown in Table 3.

The results of precision and accuracy experiments at four different concentrations are shown in Table 4. Intra-day accuracy (RE%) of the six analytes ranged from 0.0 to 11.7% while inter-day accuracy ranged from 0.7 to 7.8%. The intra- and inter-day precision (RSD%) varied from 1.7 to 14.5%, and 4.2 to 13.8%, respectively. Overall, results indicated that the method was accurate and precise at detecting each pesticide.

The stability results for the six pesticides are summarized in Table 5. Atrazine, aceto chlor, chlorpyrifos, α -endosulfan, and β -endosulfan were stable in human plasma when left on the bench top for 2 h before processing, in the autosampler for 24 h after processing, and after storage at -20°C for 17 days and two freeze-thaw cycles. Dichlorvos was unstable in human plasma when left on the bench top for 2 h before processing, and a degradation of about 50% was noted. It was also unstable after storage at -20°C for 17 days and two freeze-thaw cycles, and a degradation of about 20% was noted. However, dichlorvos was stable in the autosampler for 24 h after processing. This indicates that some endogenous components of plasma may

Table 3. The regression equation, correlation coefficient, and RSDs for LLOQ

Pesticide	Regression equation	Correlation coefficient (r^2)	RSDs for LLOQ (%)
Dichlorvos	$y = 0.6355x - 7.556e - 003$	0.9956	5.2
Atrazine	$y = 0.3016x - 9.134e - 004$	0.9964	8.4
Acetochlor	$y = 0.1769x + 1.095e - 003$	0.9912	8.3
Chlorpyrifos	$y = 0.1018x - 1.193e - 003$	0.9952	7.5
α -Endosulfan	$y = 0.03923x + 5.972e - 003$	0.9953	9.5
β -Endosulfan	$y = 0.03969x + 1.938e - 003$	0.9960	9.8

Table 4. The precision and accuracy with which six pesticides were detected

Pesticide	Concentration ($\mu\text{g/mL}$)	Intra-day		Inter-day	
		Accuracy (RE%)	Precision (RSD%)	Accuracy (RE%)	Precision (RSD%)
Dichlorvos	0.05	0.0	5.7	2.0	9.0
	0.15	-3.3	4.0	2.0	5.7
	0.8	7.4	2.2	6.0	6.2
Atrazine	8	-11.7	3.7	-4.5	7.4
	0.05	-2.0	7.8	-2.0	8.1
	0.15	-7.3	2.4	-2.7	5.1
Acetochlor	0.8	2.1	2.7	3.7	4.7
	8	-10.7	4.4	-2.2	7.7
	0.05	-10.0	7.6	-6.0	10.4
Chlorpyrifos	0.15	-6.0	4.4	-4.7	6.6
	0.8	5.9	8.6	3.6	5.4
	8	9.4	4.0	1.7	7.5
α -Endosulfan	0.05	-6.0	5.3	-2.0	9.2
	0.15	-8.7	4.4	-5.3	7.2
	0.8	0.3	7.5	0.8	4.8
β -Endosulfan	8	6.2	5.0	-3.6	8.6
	0.05	-4.0	14.5	-4.0	13.8
	0.15	0.7	4.6	-0.7	5.1
Chlorpyrifos	0.8	5.3	6.8	3.4	4.4
	8	-1.1	4.4	-7.8	7.1
	0.05	8.0	6.0	-2.0	12.1
Atrazine	0.15	6.7	4.4	-1.3	8.2
	0.8	-1.0	1.7	1.3	4.2
	8	-5.0	2.2	-5.2	8.3

Table 5. The stability (RE %) of the six pesticides in human plasma under various conditions

Pesticide	Concentration ($\mu\text{g/mL}$)	Left on bench top (2 h)	Left in autosampler (24 h)	Two freeze-thaw cycles	17 days at -20°C
Dichlorvos	0.15	-54.0	0.6	-24.7	-21.3
	8	-47.4	-10.2	-18.3	-15.6
Atrazine	0.15	-0.7	13.3	3.3	13.3
	8	5.5	0.04	4.7	15.0
Acetochlor	0.15	-0.7	5.3	4.7	13.3
	8	10.9	-2.5	13.2	12.6
Chlorpyrifos	0.15	-11.3	4.7	-12.0	11.3
	8	7.6	-3.5	10.1	11.9
α -Endosulfan	0.15	-8.7	14.0	3.3	13.3
	8	-6.3	-6.3	-7.1	7.4
β -Endosulfan	0.15	-10.0	6.0	-4.0	11.3
	8	7.3	-1.4	-1.7	6.8

affect the stability of dichlorvos, so it should be detected immediately after collecting blood from a patient.

3.4. Application

This method was used to determine the presence of pesticides in 35 plasma samples from 15 patients with acute self-poisoning seen in the Emergency Department

of Qilu Hospital of Shandong University.

The plasma was collected before and after treatment. Dichlorvos was found in 13 plasma samples at a concentration ranging from 0.05-0.725 $\mu\text{g/mL}$ with a mean concentration of 0.289 $\mu\text{g/mL}$. Atrazine was found in six plasma samples from two patients before and after treatment at a concentration ranging from 0.070-0.240 $\mu\text{g/mL}$ with a mean concentration of 0.261

µg/mL. Acetochlor was found in one plasma sample at a concentration of 0.153 µg/mL. The determined concentrations of pesticides in samples obtained from patients with acute self-poisoning are shown in Table 6. The total ion current chromatograms of patients with acute self-poisoning are shown in Figure 2. In addition, another peak (marked with “x”) appeared near the retention time of dichlorvos in eight plasma samples obtained from patients with suspected dichlorvos poisoning, as shown in Figure 2B. This peak has the same ions of m/z 109 and 79 but no ion of m/z 185. This may be a degradation product or metabolite of dichlorvos and needs to be studied further.

Table 6. The determined concentrations of samples from patients with acute self-poisoning

Samples	Concentration (µg/mL)		
	dichlorvos	atrazine	acetochlor
1	0.725	0.240	0.153
2	0.700	0.180	-
3	0.270	0.700	-
4	0.050	0.277	-
5	0.185	0.091	-
6	0.050	0.079	-
7	0.356	-	-
8	0.120	-	-
9	0.193	-	-
10	0.558	-	-
11	0.321	-	-
12	0.150	-	-
13	0.075	-	-
Mean	0.289	0.261	-

Note: The samples were obtained from different patients poisoned with dichlorvos, atrazine, or acetochlor.

Some methods such as GC-NPD, GC-MS, and liquid chromatography-mass spectrometry (LC-MS) have been used to determine the presence of dichlorvos (21), acetochlor (22), atrazine (23), chlorpyrifos (24), and α-endosulfan and β-endosulfan (25) in plasma or serum. However, these methods can only detect one or two pesticides, and liquid-liquid extraction or SPE requires a somewhat complicated method of sample preparation. In this study, a method of GC-MS was developed and fully validated for simultaneous determination of the presence of the six pesticides mentioned above. It is simple since samples are prepared using PPT, economical since acetonitrile is used as the protein precipitator, and sensitive with an LLOQ of 0.05 µg/mL. To the extent known, this is the first study to report a method for simultaneous determination of the presence of six pesticides in patients with acute poisoning.

4. Conclusion

In this study, a GC-MS method was developed to simultaneously quantify the concentrations of six pesticides including dichlorvos, atrazine, acetochlor, chlorpyrifos, α-endosulfan, and β-endosulfan in human plasma. The method is both sensitive and selective, and it is readily validated. Sample preparation is accomplished through convenient and economical protein precipitation and only requires 200 µL of plasma. This method can be used for routine clinical detection of dichlorvos, atrazine, acetochlor, chlorpyrifos, α-endosulfan, and β-endosulfan in patients with acute self-poisoning.

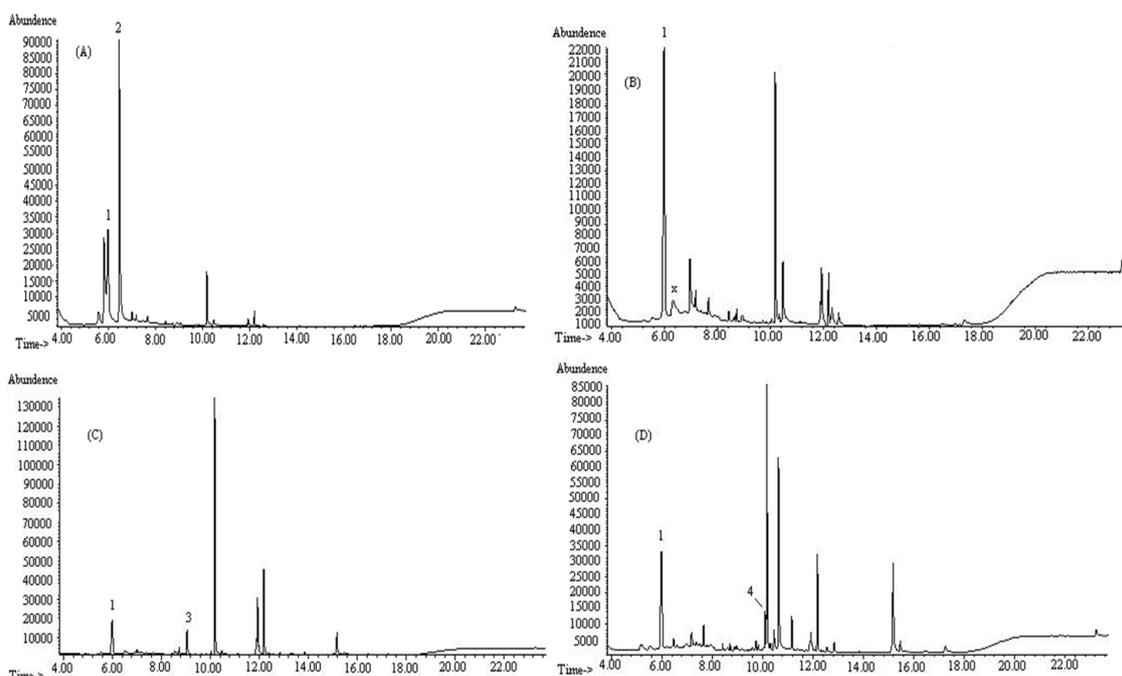


Figure 2. The total ion current chromatograms of dichlorvos (A), suspected dichlorvos (B), atrazine (C), and acetochlor (D) in patients with acute poisoning. (1-Borneol; 2-Dichlorvos; 3-Atrazine; 4-Acetochlor; x-uncertain compound).

References

- González-Alzaga B, Lacasaña M, Aguilar-Garduño C, Rodríguez-Barranco M, Ballester F, Rebagliato M, Hernández AF. A systematic review of neurodevelopmental effects of prenatal and postnatal organophosphate pesticide exposure. *Toxicol Lett.* 2014; 230:104-121.
- Kumar L, Agarwal SS, Chavali KH, Mestri SC. Homicide by organophosphorus compound poisoning: A case report. *Med Sci Law.* 2009; 49:136-138.
- Koutros S, Langseth H, Grimsrud TK, Barr DB, Vermeulen R, Portengen L, Wacholder S, Freeman LE, Blair A, Hayes RB, Rothman N, Engel LS. Prediagnostic serum organochlorine concentrations and metastatic prostate cancer: A nested case-control study in the Norwegian janus serum bank cohort. *Environ Health Perspect.* 2015; 123:867-872.
- Chaparro-Narváez P, Castañeda-Orjuela C. Mortality due to pesticide poisoning in Colombia, 1998-2011. *Biomedica.* 2015; 35:90-102.
- Chang SS, Lu TH, Sterne JA, Eddleston M, Lin JJ, Gunnell D. The impact of pesticide suicide on the geographic distribution of suicide in Taiwan: A spatial analysis. *BMC Public Health.* 2012; 12:260-270.
- Barr DB, Barr JR, Driskell WJ, Hill RH Jr, Ashley DL, Needham LL, Head SL, Sampson EJ. Strategies for biological monitoring of exposure for contemporary-use pesticides. *Toxicol Ind Health.* 1999; 15:168-179.
- Barr DB, Barr JR, Maggio VL, Whitehead RD Jr, Sadowski MA, Whyatt RM, Needham LL. A multi-analyte method for the quantification of contemporary pesticides in human serum and plasma using high-resolution mass spectrometry. *J Chromatogr B Analyt Technol Biomed Life Sci.* 2002; 778:99-111.
- Kapka-Skrzypczak L, Cyranka M, Skrzypczak M, Kruszewski M. Biomonitoring and biomarkers of organophosphate pesticides exposure-State of the art. *Ann Agric Environ Med.* 2011; 18:294-303.
- Leyya-Morales JB, Valdez-Torres JB, Bastidas-Bastidas PJ, Betancourt-Lozano M. Validation and application of a multi-residue method, using accelerated solvent extraction followed by gas chromatography, for pesticides quantification in soil. *J Chromatogr Sci.* 2015; 53:1623-1630.
- Zayats MF, Leschev SM, Zayats MA. A novel method for the determination of some pesticides in vegetable oils based on dissociation extraction followed by gas chromatography-mass spectrometry. *Food Addit Contam Part A Chem Anal Control Expo Risk Assess.* 2016; 33:1337-1345.
- Mondal R, Kole RK, Bhattacharyya A. Validation of multiresidue method for analysis of 31 pesticides in rice using gas chromatography-tandem mass spectrometry. *J AOAC Int.* 2017; 100:1094-1101.
- Li L, Zhou Y. Detection of carbamate pesticide in plasma by SPE combined with GC-FID and GC-MS. *Chin J Pharmaceu Anal.* 2000; 20:154-157.
- Liu J, Suzukib O. Conditions of solid-phase extraction for the mixture of organophosphates and synthetic pyrethroids in human body fluids. *Forensic Sci Int.* 1999; 99:159-161.
- Paqliuca G, Gazzotti T, Zironi E, Sticca P. Residue analysis of organophosphorus pesticides in animal matrices by dual column capillary gas chromatography with nitrogen-phosphorus detection. *J Chromatogr A.* 2005; 1071:67-70.
- Pavlic M, Haidekker A, Grubwieser P, Rabl W. Fatal intoxication with omethoate. *Int J Legal Med.* 2002; 116:238-241.
- Naksen W, Prapamontol T, Mangklabruks A, Chantara S, Thavornnutikarn P, Robson MG, Ryan PB, Barr DB, Panuwet P. A single method for detecting 11 organophosphate pesticides in human plasma and breastmilk using GC-FPD. *J Chromatogr B Analyt Technol Biomed Life Sci.* 2016; 1025:92-104.
- Wang Y, Du R. Simultaneous extraction of trace organophosphorus pesticides from plasma sample by automated solid phase extraction and determination by gas chromatography coupled with pulsed flame photometric detector. *Forensic Sci Int.* 2010; 198:70-73.
- Kim M, Song NR, Hong J, Lee J, Pyo H. Quantitative analysis of organochlorine pesticides in human serum using headspace solid-phase microextraction coupled with gas chromatography-mass spectrometry. *Chemosphere.* 2013; 92:279-285.
- Bedi JS, Gill JP, Kaur P, Sharma A, Aulakh RS. Evaluation of pesticide residues in human blood samples from Punjab (India). *Vet World.* 2015; 8:66-71.
- Wang D, Behniwal P, Fan R, Simon Ip HS, She J. Matrix effects in analysis of dialkyl phosphate metabolites of organophosphate pesticides in urine by gas chromatography/tandem mass spectrometer. *J Environ Sci Health B.* 2013; 48:177-182.
- Sun ZJ, DU SM, Yang CB. Method for determining concentration of dichlorvos in serum by gas chromatography. *Zhonghua Lao Dong Wei Sheng Zhi Ye Bing Za Zhi.* 2013; 31:69-71. (in Chinese)
- Zhao JJ, DU SM, Zhang YX. Determination of acetochlor in human serum with gas chromatography. *Zhonghua Lao Dong Wei Sheng Zhi Ye Bing Za Zhi.* 2010; 28:620-621. (in Chinese)
- Brzezicki JM, Andersen ME, Cranmer BK, Tessari JD. Quantitative identification of atrazine and its chlorinated metabolites in plasma. *J Anal Toxicol.* 2003; 27:569-573.
- Liao HT, Hsieh CJ, Chiang SY, Lin MH, Chen PC, Wu KY. Simultaneous analysis of chlorpyrifos and cypermethrin in cord blood plasma by online solid-phase extraction coupled with liquid chromatography-heated electrospray ionization tandem mass spectrometry. *J Chromatogr B Analyt Technol Biomed Life Sci.* 2011; 879:1961-1966.
- Chan MP, Mohd MA. Analysis of endosulfan and its metabolites in rat plasma and selected tissue samples by gas chromatography-mass spectrometry. *Environ Toxicol.* 2005; 20:45-52.

(Received March 19, 2018; Revised April 11, 2018; Accepted April 25, 2018)

Performing laparoscopic surgery – Perspectives of young Chinese hepatobiliary surgeons

Xin Zhao¹, Peipei Song^{2,*}, Yuhua Zhang³, Jiwei Huang⁴

¹Division of General Surgery, 302 Military Hospital of China, Beijing, China;

²Department of Digestive Surgery, Nihon University School of Medicine, Tokyo, Japan;

³Department of Hepatobiliary, Pancreatic and Minimally Invasive Surgery, Zhejiang Provincial People's Hospital, Hangzhou, China;

⁴Department of Liver Surgery, Division of Liver Transplantation, West China Hospital, Sichuan University, Chengdu, China.

Summary Laparoscopic liver resection (LLR) has garnered attention as a new form of liver surgery. In China, many hepatobiliary surgeons are now encouraging the examination and assessment of LLC in order to improve its outcomes, and several young hepatobiliary surgeons recently shared their clinical experiences and the results of their research in presentations at the Akamon Forum as part of the 118th Annual Congress of the Japan Surgical Society, which was held April 5-7, 2018 in Tokyo, Japan. In China, LLR has gradually improved over the past 20 years, including both expanded indications and improved surgical approaches. However, China is a vast country, and the level of medical care varies nationwide. Medical facilities that can perform advanced laparoscopic techniques are currently limited to those in large cities. Moreover, additional clinical studies of the long-term oncological outcomes of LLR need to be performed in the future.

Keywords: Laparoscopic liver resection, laparoscopic anatomical liver resection, ICG fluorescence imaging

Since its introduction in the late 1980s, various forms of abdominal surgery have been performed laparoscopically. Laparoscopic liver resection (LLR) has garnered attention as a new form of liver surgery, and evidence of its effectiveness is being assembled and reliable guidelines are being drafted (1-3). Many hepatobiliary surgeons are now encouraging the examination and assessment of LLC in order to improve its outcomes (4-6). Recently, several young Chinese hepatobiliary surgeons shared their clinical experiences and the results of their research at the 118th Annual Congress of the Japan Surgical Society, which was held April 5-7, 2018 in Tokyo, Japan (7).

With the support of Professor Norihiro Kokudo, the Congress President of the 118th Annual Congress of Japan Surgical Society, a special session entitled the Akamon Forum was held on April 6. More than

twenty Chinese hepatobiliary surgeons from leading hepato-pancreato-biliary centers in China (including surgeons from Tsinghua Changgung Hospital, West China Hospital of Sichuan University, Hunan Provincial People's Hospital, The Third Affiliated Hospital of Sun Yat-sen University, Peking University Cancer Hospital, 302 Military Hospital of China, Sun Yat-sen Memorial Hospital of Sun Yat-sen University, Beijing Shijitan Hospital of Capital Medical University, and Zhejiang Provincial People's Hospital) attended this special forum to report the results of their research into liver tumors, cholangiocarcinoma, pancreatic neoplasms, pediatric liver transplantation, and portal hypertension.

Professor Kokudo delivered warm opening remarks. Since all of the young Chinese hepatobiliary surgeons attending the Akamon Forum had studied in Japan, Professor Kokudo welcomed them back to Tokyo to share their clinical experiences and the results of their research. Professor Kokudo stressed how the Forum would surely enhance communication between young surgeons from Japan and China. Professor Kiyoshi Hasegawa from the University of Tokyo and Professor Wei Tang from the National Center for Global Health

*Address correspondence to:

Dr. Peipei Song, Department of Digestive Surgery, Nihon University School of Medicine, 30-1 Oyaguchikamimachi, Itabashi-ku, Tokyo 173-8610, Japan.

E-mail: ppsong-tky@umin.ac.jp

and Medicine delivered closing remarks. The professors remarked on how the Forum represented an exceptional platform for academic exchanges and they encouraged the Forum to be held in the future.

In line with the theme of the 118th Annual Congress of the Japan Surgical Society – "Looking for New Findings in Surgery" – the topics of laparoscopic anatomical liver resection, ICG fluorescence imaging in LLR, and laparoscopic resection of hilar cholangiocarcinoma led to an interesting discussion among the audience at the Akamon Forum.

LLR was first reported in the early 1990s worldwide (8,9). The Eastern Hepatobiliary Surgery Hospital performed LLR for the first time in China in 1994 (10). After more than 20 years, LLR has gradually improved, and the total number of surgical cases has rapidly increased. Surgical indications have gradually expanded. Laparoscopic resection for recurrent liver cancer and laparoscopic hepatectomy for a living donor in living donor liver transplantation have also been reported by some facilities.

In addition to expanded indications, surgical approaches to LLR have gradually progressed from wedge resection and irregular resection to laparoscopic regular resection, hemi-hepatectomy, and anatomical segmentectomy. As laparoscopic techniques and equipment continue to improve, a variety of energy devices such as ultrasonic scalpels, argon bean coagulators, bipolar coagulators, and CUSA have become routine instruments. Several studies have reported on the safety and feasibility of LLR (11,12). According to published studies (11-14), LLR has several advantages: less blood loss, a lower rate of complications, a shorter duration of hospitalization, and a lower incidence of repeat liver resection and with no difference in oncological outcomes for liver malignancies compared to open liver resection. Some high-volume medical facilities in China have gradually begun to perform LLR more often, and the rate at which LLR is performed has even exceeded that of open liver resection. Moreover, some primary hospitals in China are performing LLR as well. However, more evidence needs to be compiled in the future.

In recent years, increased use of ultrasound in laparoscopic surgery has resulted in heightened clinical research into laparoscopic anatomical liver resection. Laparoscopic anatomical hepatectomy through exposing the hepatic veins and application of methylene blue dye injection in laparoscopic hepatectomy have been successfully performed. The recent use of ICG fluorescence imaging-guided hepatectomy in LLR has become a topic of considerable interest. A search of the literature revealed that several facilities in China have performed laparoscopic anatomical liver resection in more than 50 cases, respectively (15-17).

As patients worldwide increasingly opt for less invasive procedures, laparoscopic surgery will gain

even greater prominence. As the presentations at the Akamon Forum illustrate, young Chinese hepatobiliary surgeons are increasingly turning their attention to laparoscopic procedures. However, China is a vast country, and the level of medical care varies nationwide. Medical facilities that can perform advanced laparoscopic techniques are currently limited to those in large cities. Moreover, additional clinical studies of the long-term oncological outcomes of LLR need to be performed in the future.

References

- Cheung TT, Han HS, She WH, Chen KH, Chow PKH, Yoong BK, Lee KF, Kubo S, Tang CN, Wakabayashi G. The Asia Pacific consensus statement on laparoscopic liver resection for hepatocellular carcinoma: A report from the 7th Asia-Pacific Primary Liver Cancer Expert Meeting held in Hong Kong. *Liver Cancer*. 2018; 7:28-39.
- Buell JF, Cherqui D, Geller DA, *et al*. The international position on laparoscopic liver surgery: The Louisville Statement, 2008. *Ann Surg*. 2009; 250:825-830.
- Wakabayashi G, Cherqui D, Geller DA, *et al*. Recommendations for laparoscopic liver resection: A report from the second international consensus conference held in Morioka. *Ann Surg*. 2015; 261:619-629.
- Chen K, Pan Y, Zhang B, Liu XL, Maher H, Zheng XY. Laparoscopic versus open surgery for hepatocellular carcinoma: A meta-analysis of high-quality case-matched studies. *Can J Gastroenterol Hepatol*. 2018; 2018:1746895.
- Chen J, Bai T, Zhang Y, Xie ZB, Wang XB, Wu FX, Li LQ. The safety and efficacy of laparoscopic and open hepatectomy in hepatocellular carcinoma patients with liver cirrhosis: A systematic review. *Int J Clin Exp Med*. 2015; 8:20679-20689.
- Morise Z, Ciria R, Cherqui D, Chen KH, Belli G, Wakabayashi G. Can we expand the indications for laparoscopic liver resection? A systematic review and meta-analysis of laparoscopic liver resection for patients with hepatocellular carcinoma and chronic liver disease. *J Hepatobiliary Pancreat Sci*. 2015; 22:342-352.
- Japan Surgical Society. The 118th Annual Congress of the Japan Surgical Society. <https://www.jssoc.or.jp/jss118/index.html> (accessed on April 6, 2018)
- Reich H, McGlynn F, DeCaprio J, Budin R. Laparoscopic excision of benign liver lesions. *Obstet Gynecol*. 1991; 78:956-958.
- Gagner M, Rheault M, Dubuc J. Laparoscopic partial hepatectomy for liver tumor. *Surg Endosc*. 1992; 6:97-98.
- Zhou WP, Sun ZH, Wu MC, Chen H, Zhang BH, Zheng CZ, Shen YM, Qiu M. The first report of laparoscopic hepatectomy. *Journal of Hepatobiliary Surgery*. 1994; 2:82. (in Chinese)
- Tranchart H, Dagher I. Laparoscopic liver resection: A review. *J Visc Surg*. 2014; 151:107-115.
- Nguyen KT, Gamblin TC, Geller DA. World review of laparoscopic liver resection-2,804 patients. *Ann Surg*. 2009; 250:831-841.
- Martin RC, Scoggins CR, McMasters KM. Laparoscopic hepatic lobectomy: Advantages of a minimally invasive approach. *J Am Coll Surg*. 2010; 210:627-634.
- Franken C, Lau B, Putschakayala K, DiFronzo LA.

- Comparison of short-term outcomes in laparoscopic vs open hepatectomy. *JAMA Surg.* 2014; 149:941-946.
15. Cheng W, Zuo Z, Liu Y, Wu YF, Zhu SW, Peng C, Shen XB, Yin XM. Analysis of 97 cases of laparoscopic hepatic vein anatomical hepatectomy. *Chinese Journal of Practical Surgery.* 2017, 37:555-558. (in Chinese)
 16. Meng J, Du ST, Zhang JG. Comparative study of laparoscopic liver wedge resection and anatomic resection. *Journal of Laparoscopic Surgery.* 2016; 5:321-324. (in Chinese)
 17. Xu DB, Li CG, Hu MG, Zhao ZM, Liu R. Application of endoscopic linear cutting closure in laparoscopic anatomical hepatectomy. *Journal of Laparoscopic Surgery.* 2015:426-428. (in Chinese)

(Received April 9, 2018; Accepted April 24, 2018)

Strengthening maternal and child health in China: Lessons from transforming policy proposals into action

Xiaoguang Yang¹, Shenglan Tang^{2,3}, Gavin Yamey², Xu Qian^{1,4,*}

¹ School of Public Health, Fudan University, Shanghai, China;

² Duke Global Health Institute, Duke University, Durham, USA;

³ Global Health Research Center at Duke Kunshan University, Kunshan, China;

⁴ Global Health Institute, Fudan University, Shanghai, China.

Summary China has made impressive achievements in improving maternal and child health (MCH) over the past few decades. This paper uses a policy lens to examine reasons for these achievements as well as barriers to further success. We found that strong governmental commitment and leadership, effective coordination, proactive participation of different stakeholders, and the provision of adequate resources were associated with China's success in improving MCH outcomes. Other low- and middle-income countries can learn valuable lessons from China's experience. These lessons include *i*) prioritizing MCH on the national development agenda, *ii*) keeping national ownership over health development cooperation, and *iii*) establishing effective monitoring, evaluation and accountability mechanisms for MCH programs.

Keywords: Maternal and child health, policy implementation, enabling factors, China

1. Introduction

China has significantly improved its maternal and child health (MCH) outcomes since the establishment of the People's Republic of China in 1949. Data available at the national level show that the key indicators of MCH in China have continually improved since the early 1990s (Table 1). From 1991 to 2013, China's under-five mortality rate fell by 80% and its maternal mortality ratio (MMR) fell by 71% (1).

These achievements are due largely to the country's rapid, sustained economic growth, the continuing enhancement of the MCH service delivery system (2), and, as we discuss in this paper, the effective development and implementation of health policies and strategies. In this paper, we attempt to distill lessons on implementing MCH strategies by taking an overarching policy view. Such lessons are based on decades of experience with both successes and setbacks. We hope

these lessons will be useful to other low- and middle-income countries (LMICs) that still face daunting challenges in reducing their maternal and child mortality rates (3).

2. The Evolution of China's MCH strategy

China's achievement in improving MCH is based on its well-established policy framework. Figure 1 illustrates a timeline of major events and milestones in the development of China's MCH policies and initiatives since the 1980s. During the 1980s, a series of service standards, guidelines and regulations on MCH were developed and adopted, including the urban obstetric care quality standards and requirements (in 1980), the regulation of maternal and child health care (1986), and the regulation of maternal health care in rural areas (1989).

The 1990s witnessed a major shift in China's MCH strategy. In this period, the Chinese government made a number of commitments to international conventions related to the health and wellbeing of women and children, such as the Declaration on the Survival, Protection and Development of Children at the World Summit for Children (1990) and the Beijing Declaration and Action Plan at the 4th World Conference on Women

Released online in J-STAGE as advance publication April 15, 2018.

*Address correspondence to:

Prof. Xu Qian, School of Public Health, Fudan University, Box 175, 138 Yi Xue Yuan Road, Shanghai 200032, China.
E-mail: xqian@fudan.edu.cn

Table 1. Key indicators of MCH in China from 1991-2014

Indicators	1991	1995	2000	2005	2010	2014
Infant mortality rate (deaths per 1,000 live births)	50.2	36.4	32.2	19.0	13.1	8.9
Under-5 mortality rate (deaths per 1,000 live births)	61.0	44.5	39.7	22.5	16.4	11.7
Maternal mortality ratio (deaths per 100,000 live births)	80.0	61.9	53.0	47.7	30.0	21.7

Source: World Bank

(1995). These international conventions provided China with opportunities to develop a policy framework aimed at improving MCH in China. China developed a series of laws, inter-connected strategies, and action plans, particularly the Action Plan of the China National Program for Children's Development (1992) and the China National Program for Women's Development (1995) (known as "Two Programs" in short). The Chinese government actively cooperated with international agencies to implement programs aimed at improving MCH status. These efforts established a strong foundation for the development of MCH in the 21st century, particularly the much improved MCH service delivery system, which covers both urban and rural areas. The government introduced multiple strategies to reduce the financial burden for institution-based childbirth, such as service price regulation, medical aid for the poor, reimbursement by health insurance, and direct subsidies (4).

3. The key factors enabling policy implementation

China's achievements in MCH are due in large part to defined policies and forceful implementation. Below we focus on several key determinants for successful policy implementation which we regard as major lessons of MCH promotion in China.

3.1. Transforming political commitment into implementable policy and strategy

The Chinese government's commitment to international goals (e.g., its commitment to achieving the health related Millennium Development Goals [MDGs]) provided an excellent opportunity for China to put MCH firmly on the government agenda (5), which facilitated the initiation of domestic policies (in the form of laws, regulations, and implementation guidelines) including the most important, "Two Programs". The "Two Programs" policy was developed with clear targets, timelines, and accountability. For example, the second "Two Programs" in 2001-2010 stated that the MMR should be reduced by one quarter by 2010, using the MMR in 2000 (53/100,000) as the baseline. With this set target, various policy approaches were adopted, such as boosting technical and financial capacity to improve MCH, instituting the national "Jiang Xiao" Project to reduce MMR and eliminate neonatal tetanus, and

promoting institutional childbirth.

3.2. Setting a unified administrative system and strong accountability for policy implementation

The unified, hierarchical administrative system in China helped to ensure the effective implementation and accountability of the policies defined by the central government throughout the country. Once specific MCH targets and their indicators have been set, the different levels of government, as responsible implementation agencies, have the obligation to ensure these targets are met in a specified timeframe. Poor performance would lead to a bad evaluation for the local governors, which in turn might (indirectly) affect their further political promotion. Moreover, MCH mortality surveillance systems were employed by provinces in order to reduce under-reported maternal deaths (6).

3.3. Mobilizing key stakeholders

Ultimately, actors are the key agents for transforming defined policies into practice. For example, the local government should adapt the national MCH policy into specific, localized action plans. Under the accountability system, the local government then has the motivation to put more effort and resources into MCH-related endeavors, in order to achieve the targets set in the national action plan. Other key actors are the frontline health workers. The government's commitment to improving MCH, including its investment in health human resources, has greatly helped frontline health workers build their confidence and their willingness to conduct routine work. In addition, the strong sense of altruism shown by many MCH workers, particularly those from the remote rural areas, has also been an important factor in improving performance.

4. The valuable lessons from China's experience for other low- and middle-income countries

China has made great progress in MCH, especially in improving maternal and child survival. Nevertheless, problems of regional disparities in MCH and sub-optimal service coverage for migrant populations still need to be addressed (7). We believe that the Chinese experiences and practices of improving MCH over the past three decades may offer several key lessons for

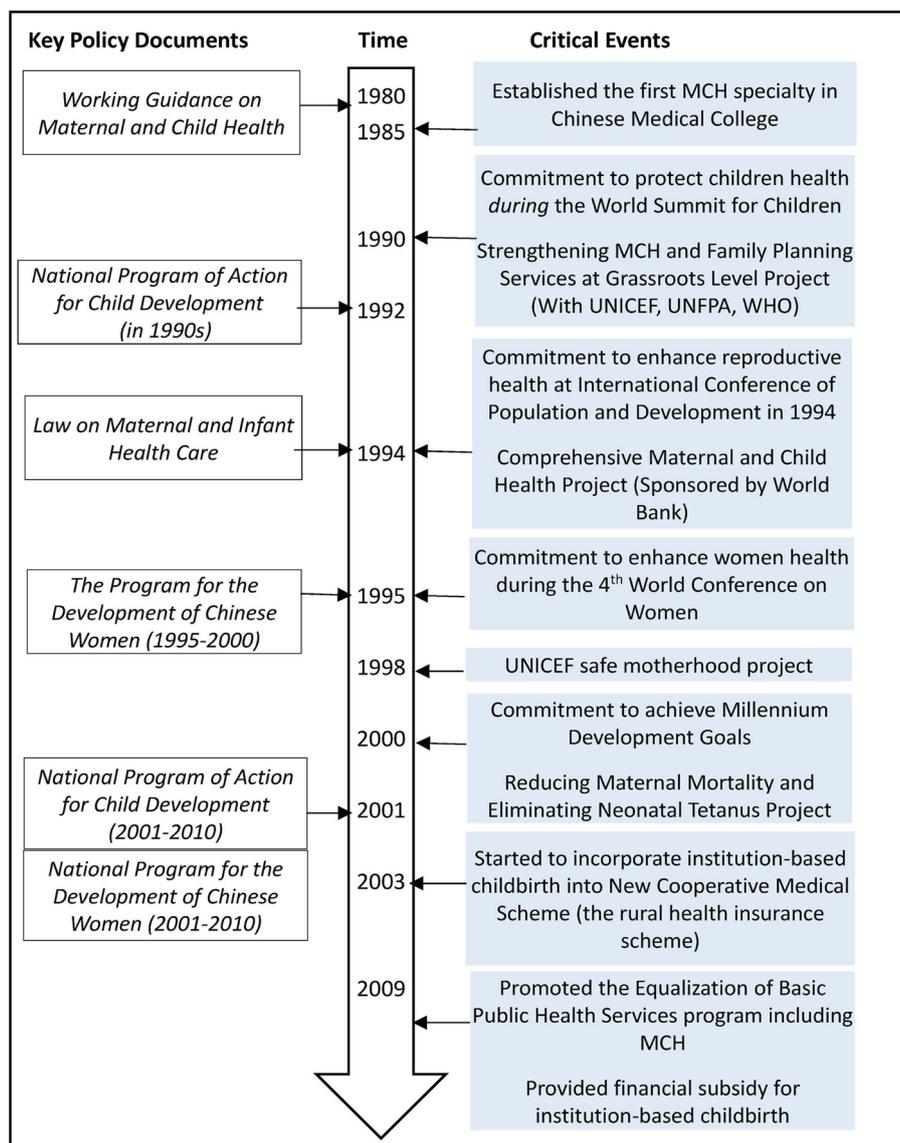


Figure 1. Brief timeline of China's MCH strategy since 1980.

other developing countries.

Government commitment to MCH The government commitment and political will to improve MCH placed it high on the national agenda and allowed MCH to be incorporated into national policy development. For example, the commitment made by the Chinese government to international conventions in the early 1990s greatly promoted the priority of MCH on the governmental agenda, compared with the 1980s. As a result, the decision makers at both national and local levels were pushed to take effective action to improve MCH. For LMICs, the Sustainable Development Goals (SDGs) provide a good opportunity to incorporate MCH into their own development strategies or governmental agendas. Such actions would give them the opportunity to access more resources for implementation, such as funding, personnel, and supplies, which MCH urgently needs.

Domestically driven approach The policy and practice of MCH in LMICs must have "national

ownership", instead of being "donor driven". Programs should be domestically driven and designed around individual countries' political and social contexts, with support from the international donors. China used to be one of the largest recipients of health development cooperation. However, as shown in Figure 1, the financial and technical support from international donors has gradually given way to domestic funding and expertise, which has enhanced the existing health service delivery system, including MCH. Moreover, with the successful implementation of the international aid program, more domestic funding was leveraged to co-finance MCH interventions in the long run. Nowadays, LMICs cooperate with many multilateral and bilateral health and development agencies, and it is necessary to make use of the external support to mobilize more domestic resources, and scale up implementation in a sustainable way.

Strong monitoring and accountability mechanism Effective health development cooperation programs

rely on strong monitoring, evaluation and accountability mechanisms. In China, the monitoring and evaluation indicators written in the "Two Programs" played a key role in enhancing the MCH policy implementation. These indicators were not only acting as a means of supervision and monitoring for frontline health workers, but also as the performance assessment for local government incentives. Furthermore, more importance should be put on health information systems in order to provide reliable data to inform decisions.

Reaching the SDGs will be a long and difficult task for many LMICs. We hope that China's experience with MCH health policy design, action plan implementation, resource allocation, as well as monitoring and accountability can provide valuable lessons to help LMICs achieve the SDGs.

Acknowledgements

The authors appreciate the sponsor from China-United Kingdom (UK) Global Health Support Programme funded by UK Department for International Development.

References

1. Kuruvilla S, Schweitzer J, Bishai D, *et al.* Success factors for reducing maternal and child mortality. Bull World Health Organ. 2014; 92:533-544.
2. World Health Organization. Success factors for women's and children's health: Policy and programme highlights from 10 fast-track countries. http://www.who.int/entity/pmnch/knowledge/publications/success_factors_highlights.pdf (accessed on February 26, 2018)
3. Liang J, Li X, Dai L, Zeng W, Li Q, Li M, Zhou R, He C, Wang Y, Zhu J. The changes in maternal mortality in 1000 counties in mid-Western China by a government-initiated intervention. PLoS One, 2012; 7:e37458.
4. Yang XG, Wang SS, Qian X, Multiple Measures to Alleviate Financial Burden of Facility-Based Childbirth, GHSP Project Policy brief, 2016 , [http://cps.moh.gov.cn/website-webapp/ewebeditor/uploadfile/E07-OP102-4\(C2\)%20Multiple%20Measures%20to%20Alleviate%20Financial%20Burden%20of%20Fa-cility-Based%20Childbirth%20%20with%20mark.pdf](http://cps.moh.gov.cn/website-webapp/ewebeditor/uploadfile/E07-OP102-4(C2)%20Multiple%20Measures%20to%20Alleviate%20Financial%20Burden%20of%20Fa-cility-Based%20Childbirth%20%20with%20mark.pdf) (accessed on Feb 26, 2018)
5. Unicef. Committing to child survival: A promise renewed. https://www.unicef.org/publications/files/APR_2015_9_Sep_15.pdf (accessed on February 26, 2018)
6. Kingdon JW. Agendas, alternatives, and public policies. Agendas, alternatives, and public policies. Little, Brown, Boston, USA, 1984; pp. 165-169.
7. Li Y, Zhang Y, Fang S, Liu S, Liu X, Li M, Liang H, Fu H. Analysis of inequality in maternal and child health outcomes and mortality from 2000 to 2013 in China. Int J Equity Health. 2017; 16:66.

(Received March 2, 2018; Revised April 7, 2018; Accepted April 8, 2018)

Guide for Authors

1. Scope of Articles

BioScience Trends is an international peer-reviewed journal. BioScience Trends devotes to publishing the latest and most exciting advances in scientific research. Articles cover fields of life science such as biochemistry, molecular biology, clinical research, public health, medical care system, and social science in order to encourage cooperation and exchange among scientists and clinical researchers.

2. Submission Types

Original Articles should be well-documented, novel, and significant to the field as a whole. An Original Article should be arranged into the following sections: Title page, Abstract, Introduction, Materials and Methods, Results, Discussion, Acknowledgments, and References. Original articles should not exceed 5,000 words in length (excluding references) and should be limited to a maximum of 50 references. Articles may contain a maximum of 10 figures and/or tables.

Brief Reports definitively documenting either experimental results or informative clinical observations will be considered for publication in this category. Brief Reports are not intended for publication of incomplete or preliminary findings. Brief Reports should not exceed 3,000 words in length (excluding references) and should be limited to a maximum of 4 figures and/or tables and 30 references. A Brief Report contains the same sections as an Original Article, but the Results and Discussion sections should be combined.

Reviews should present a full and up-to-date account of recent developments within an area of research. Normally, reviews should not exceed 8,000 words in length (excluding references) and should be limited to a maximum of 100 references. Mini reviews are also accepted.

Policy Forum articles discuss research and policy issues in areas related to life science such as public health, the medical care system, and social science and may address governmental issues at district, national, and international levels of discourse. Policy Forum articles should not exceed 2,000 words in length (excluding references).

Case Reports should be detailed reports of the symptoms, signs, diagnosis, treatment, and follow-up of an individual patient. Case reports may contain a demographic profile of the patient but usually describe an unusual or novel occurrence. Unreported or unusual

side effects or adverse interactions involving medications will also be considered. Case Reports should not exceed 3,000 words in length (excluding references).

News articles should report the latest events in health sciences and medical research from around the world. News should not exceed 500 words in length.

Letters should present considered opinions in response to articles published in BioScience Trends in the last 6 months or issues of general interest. Letters should not exceed 800 words in length and may contain a maximum of 10 references.

3. Editorial Policies

Ethics: BioScience Trends requires that authors of reports of investigations in humans or animals indicate that those studies were formally approved by a relevant ethics committee or review board.

Conflict of Interest: All authors are required to disclose any actual or potential conflict of interest including financial interests or relationships with other people or organizations that might raise questions of bias in the work reported. If no conflict of interest exists for each author, please state "There is no conflict of interest to disclose".

Submission Declaration: When a manuscript is considered for submission to BioScience Trends, the authors should confirm that 1) no part of this manuscript is currently under consideration for publication elsewhere; 2) this manuscript does not contain the same information in whole or in part as manuscripts that have been published, accepted, or are under review elsewhere, except in the form of an abstract, a letter to the editor, or part of a published lecture or academic thesis; 3) authorization for publication has been obtained from the authors' employer or institution; and 4) all contributing authors have agreed to submit this manuscript.

Cover Letter: The manuscript must be accompanied by a cover letter signed by the corresponding author on behalf of all authors. The letter should indicate the basic findings of the work and their significance. The letter should also include a statement affirming that all authors concur with the submission and that the material submitted for publication has not been published previously or is not under consideration for publication elsewhere. The cover letter should be submitted in PDF format. For example of Cover Letter, please visit <http://www.biosciencetrends.com/downcentre.php> (Download Centre).

Copyright: A signed JOURNAL PUBLISHING AGREEMENT (JPA) form must be provided by post, fax, or as a scanned file before acceptance of the article. Only forms with a hand-written signature are accepted. This copyright will ensure the widest possible dissemination of information. A form facilitating transfer of copyright can be downloaded by clicking the

appropriate link and can be returned to the e-mail address or fax number noted on the form (Please visit [Download Centre](#)). Please note that your manuscript will not proceed to the next step in publication until the JPA Form is received. In addition, if excerpts from other copyrighted works are included, the author(s) must obtain written permission from the copyright owners and credit the source(s) in the article.

Suggested Reviewers: A list of up to 3 reviewers who are qualified to assess the scientific merit of the study is welcomed. Reviewer information including names, affiliations, addresses, and e-mail should be provided at the same time the manuscript is submitted online. Please do not suggest reviewers with known conflicts of interest, including participants or anyone with a stake in the proposed research; anyone from the same institution; former students, advisors, or research collaborators (within the last three years); or close personal contacts. Please note that the Editor-in-Chief may accept one or more of the proposed reviewers or may request a review by other qualified persons.

Language Editing: Manuscripts prepared by authors whose native language is not English should have their work proofread by a native English speaker before submission. If not, this might delay the publication of your manuscript in BioScience Trends.

The Editing Support Organization can provide English proofreading, Japanese-English translation, and Chinese-English translation services to authors who want to publish in BioScience Trends and need assistance before submitting a manuscript. Authors can visit this organization directly at <http://www.iacmhr.com/iac-eso/support.php?lang=en>. IAC-ESO was established to facilitate manuscript preparation by researchers whose native language is not English and to help edit works intended for international academic journals.

4. Manuscript Preparation

Manuscripts should be written in clear, grammatically correct English and submitted as a Microsoft Word file in a single-column format. Manuscripts must be paginated and typed in 12-point Times New Roman font with 24-point line spacing. Please do not embed figures in the text. Abbreviations should be used as little as possible and should be explained at first mention unless the term is a well-known abbreviation (e.g. DNA). Single words should not be abbreviated.

Title Page: The title page must include 1) the title of the paper (Please note the title should be short, informative, and contain the major key words); 2) full name(s) and affiliation(s) of the author(s), 3) abbreviated names of the author(s), 4) full name, mailing address, telephone/fax numbers, and e-mail address of the corresponding author; and 5) conflicts of interest (if you have an actual or potential conflict of interest to disclose, it must be included as a footnote on the title page of the manuscript; if no conflict of

interest exists for each author, please state "There is no conflict of interest to disclose"). Please visit [Download Centre](#) and refer to the title page of the manuscript sample.

Abstract: The abstract should briefly state the purpose of the study, methods, main findings, and conclusions. For article types including Original Article, Brief Report, Review, Policy Forum, and Case Report, a one-paragraph abstract consisting of no more than 250 words must be included in the manuscript. For News and Letters, a brief summary of main content in 150 words or fewer should be included in the manuscript. Abbreviations must be kept to a minimum and non-standard abbreviations explained in brackets at first mention. References should be avoided in the abstract. Key words or phrases that do not occur in the title should be included in the Abstract page.

Introduction: The introduction should be a concise statement of the basis for the study and its scientific context.

Materials and Methods: The description should be brief but with sufficient detail to enable others to reproduce the experiments. Procedures that have been published previously should not be described in detail but appropriate references should simply be cited. Only new and significant modifications of previously published procedures require complete description. Names of products and manufacturers with their locations (city and state/country) should be given and sources of animals and cell lines should always be indicated. All clinical investigations must have been conducted in accordance with Declaration of Helsinki principles. All human and animal studies must have been approved by the appropriate institutional review board(s) and a specific declaration of approval must be made within this section.

Results: The description of the experimental results should be succinct but in sufficient detail to allow the experiments to be analyzed and interpreted by an independent reader. If necessary, subheadings may be used for an orderly presentation. All figures and tables must be referred to in the text.

Discussion: The data should be interpreted concisely without repeating material already presented in the Results section. Speculation is permissible, but it must be well-founded, and discussion of the wider implications of the findings is encouraged. Conclusions derived from the study should be included in this section.

Acknowledgments: All funding sources should be credited in the Acknowledgments section. In addition, people who contributed to the work but who do not meet the criteria for authors should be listed along with their contributions.

References: References should be numbered in the order in which they appear in the text. Citing of unpublished results, personal communications, conference abstracts, and theses in the reference list is not recommended but these sources may be mentioned in the text. In the reference list,

cite the names of all authors when there are fifteen or fewer authors; if there are sixteen or more authors, list the first three followed by *et al.* Names of journals should be abbreviated in the style used in PubMed. Authors are responsible for the accuracy of the references. Examples are given below:

Example 1 (Sample journal reference): Inagaki Y, Tang W, Zhang L, Du GH, Xu WF, Kokudo N. Novel aminopeptidase N (APN/CD13) inhibitor 24F can suppress invasion of hepatocellular carcinoma cells as well as angiogenesis. *Biosci Trends*. 2010; 4:56-60.

Example 2 (Sample journal reference with more than 15 authors): Darby S, Hill D, Auvinen A, *et al.* Radon in homes and risk of lung cancer: Collaborative analysis of individual data from 13 European case-control studies. *BMJ*. 2005; 330:223.

Example 3 (Sample book reference): Shalev AY. Post-traumatic stress disorder: diagnosis, history and life course. In: Post-traumatic Stress Disorder, Diagnosis, Management and Treatment (Nutt DJ, Davidson JR, Zohar J, eds.). Martin Dunitz, London, UK, 2000; pp. 1-15.

Example 4 (Sample web page reference): Ministry of Health, Labour and Welfare of Japan. Dietary reference intakes for Japanese. <http://www.mhlw.go.jp/houdou/2004/11/h1122-2a.html> (accessed June 14, 2010).

Tables: All tables should be prepared in Microsoft Word or Excel and should be arranged at the end of the manuscript after the References section. Please note that tables should not be image format. All tables should have a concise title and should be numbered consecutively with Arabic numerals. If necessary, additional information should be given below the table.

Figure Legend: The figure legend should be typed on a separate page of the main manuscript and should include a short title and explanation. The legend should be concise but comprehensive and should be understood without referring to the text. Symbols used in figures must be explained.

Figure Preparation: All figures should be clear and cited in numerical order in the text. Figures must fit a one- or two-column format on the journal page: 8.3 cm (3.3 in.) wide for a single column, 17.3 cm (6.8 in.) wide for a double column; maximum height: 24.0 cm (9.5 in.). Please make sure that the symbols and numbers appeared in the figures should be clear. Please make sure that artwork files are in an acceptable format (TIFF or JPEG) at minimum resolution (600 dpi for illustrations, graphs, and annotated artwork, and 300 dpi for micrographs and photographs). Please provide all figures as separate files. Please note that low-resolution images are one of the leading causes of article resubmission and schedule delays. All color figures will be reproduced in full color in the online edition of the journal at no cost to authors.

Units and Symbols: Units and symbols

conforming to the International System of Units (SI) should be used for physicochemical quantities. Solidus notation (*e.g.* mg/kg, mg/mL, mol/mm²/min) should be used. Please refer to the SI Guide www.bipm.org/en/si/ for standard units.

Supplemental data: Supplemental data might be useful for supporting and enhancing your scientific research and BioScience Trends accepts the submission of these materials which will be only published online alongside the electronic version of your article. Supplemental files (figures, tables, and other text materials) should be prepared according to the above guidelines, numbered in Arabic numerals (*e.g.*, Figure S1, Figure S2, and Table S1, Table S2) and referred to in the text. All figures and tables should have titles and legends. All figure legends, tables and supplemental text materials should be placed at the end of the paper. Please note all of these supplemental data should be provided at the time of initial submission and note that the editors reserve the right to limit the size and length of Supplemental Data.

5. Submission Checklist

The Submission Checklist will be useful during the final checking of a manuscript prior to sending it to BioScience Trends for review. Please visit [Download Centre](#) and download the Submission Checklist file.

6. Online Submission

Manuscripts should be submitted to BioScience Trends online at <http://www.biosciencetrends.com>. The manuscript file should be smaller than 5 MB in size. If for any reason you are unable to submit a file online, please contact the Editorial Office by e-mail at office@biosciencetrends.com.

7. Accepted Manuscripts

Proofs: Galley proofs in PDF format will be sent to the corresponding author via e-mail. Corrections must be returned to the editor (proof-editing@biosciencetrends.com) within 3 working days.

Offprints: Authors will be provided with electronic offprints of their article. Paper offprints can be ordered at prices quoted on the order form that accompanies the proofs.

Page Charge: Page charges will be levied on all manuscripts accepted for publication in BioScience Trends (\$140 per page for black white pages; \$340 per page for color pages). Under exceptional circumstances, the author(s) may apply to the editorial office for a waiver of the publication charges at the time of submission.

(Revised February 2013)

Editorial and Head Office:

Pearl City Koishikawa 603
2-4-5 Kasuga, Bunkyo-ku
Tokyo 112-0003 Japan
Tel: +81-3-5840-8764
Fax: +81-3-5840-8765
E-mail: office@biosciencetrends.com

JOURNAL PUBLISHING AGREEMENT (JPA)

Manuscript No.:

Title:

Corresponding Author:

The International Advancement Center for Medicine & Health Research Co., Ltd. (IACMHR Co., Ltd.) is pleased to accept the above article for publication in BioScience Trends. The International Research and Cooperation Association for Bio & Socio-Sciences Advancement (IRCA-BSSA) reserves all rights to the published article. Your written acceptance of this JOURNAL PUBLISHING AGREEMENT is required before the article can be published. Please read this form carefully and sign it if you agree to its terms. The signed JOURNAL PUBLISHING AGREEMENT should be sent to the BioScience Trends office (Pearl City Koishikawa 603, 2-4-5 Kasuga, Bunkyo-ku, Tokyo 112-0003, Japan; E-mail: office@biosciencetrends.com; Tel: +81-3-5840-8764; Fax: +81-3-5840-8765).

1. Authorship Criteria

As the corresponding author, I certify on behalf of all of the authors that:

- 1) The article is an original work and does not involve fraud, fabrication, or plagiarism.
- 2) The article has not been published previously and is not currently under consideration for publication elsewhere. If accepted by BioScience Trends, the article will not be submitted for publication to any other journal.
- 3) The article contains no libelous or other unlawful statements and does not contain any materials that infringes upon individual privacy or proprietary rights or any statutory copyright.
- 4) I have obtained written permission from copyright owners for any excerpts from copyrighted works that are included and have credited the sources in my article.
- 5) All authors have made significant contributions to the study including the conception and design of this work, the analysis of the data, and the writing of the manuscript.
- 6) All authors have reviewed this manuscript and take responsibility for its content and approve its publication.
- 7) I have informed all of the authors of the terms of this publishing agreement and I am signing on their behalf as their agent.

2. Copyright Transfer Agreement

I hereby assign and transfer to IACMHR Co., Ltd. all exclusive rights of copyright ownership to the above work in the journal BioScience Trends, including but not limited to the right 1) to publish, republish, derivate, distribute, transmit, sell, and otherwise use the work and other related material worldwide, in whole or in part, in all languages, in electronic, printed, or any other forms of media now known or hereafter developed and the right 2) to authorize or license third parties to do any of the above.

I understand that these exclusive rights will become the property of IACMHR Co., Ltd., from the date the article is accepted for publication in the journal BioScience Trends. I also understand that IACMHR Co., Ltd. as a copyright owner has sole authority to license and permit reproductions of the article.

I understand that except for copyright, other proprietary rights related to the Work (e.g. patent or other rights to any process or procedure) shall be retained by the authors. To reproduce any text, figures, tables, or illustrations from this Work in future works of their own, the authors must obtain written permission from IACMHR Co., Ltd.; such permission cannot be unreasonably withheld by IACMHR Co., Ltd.

3. Conflict of Interest Disclosure

I confirm that all funding sources supporting the work and all institutions or people who contributed to the work but who do not meet the criteria for authors are acknowledged. I also confirm that all commercial affiliations, stock ownership, equity interests, or patent-licensing arrangements that could be considered to pose a financial conflict of interest in connection with the article have been disclosed.

Corresponding Author's Name (Signature):

Date:

